(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
18 April 2002 (18.04.2002)

PCT

(10) International Publication Number
WO 02/30268 A2

(51) International Patent Classification[7]: A61B

(21) International Application Number: PCT/US01/32045

(22) International Filing Date: 12 October 2001 (12.10.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

09/687,576	13 October 2000 (13.10.2000)	US
09/733,742	8 December 2000 (08.12.2000)	US
09/733,288	8 December 2000 (08.12.2000)	US
60/263,957	24 January 2001 (24.01.2001)	US
60/276,888	16 March 2001 (16.03.2001)	US
60/276,791	16 March 2001 (16.03.2001)	US
60/281,922	6 April 2001 (06.04.2001)	US
60/286,214	24 April 2001 (24.04.2001)	US
09/847,046	30 April 2001 (30.04.2001)	US
60/288,589	4 May 2001 (04.05.2001)	US

(71) Applicant: **EOS BIOTECHNOLOGY, INC.** [US/US]; 225A Gateway Boulevard, South San Francisco, CA 94080-7019 (US).

(72) Inventors: **GISH, Kurt, C.**; 40 Perego Terrace #2, San Francisco, CA 94131 (US). **MACK, David, H.**; 2076 Monterey Avenue, Menlo Park, CA 94025 (US). **WILSON, Keith, E.**; 219 Jeter Street, Redwood City, CA 94062 (US). **AFAR, Daniel**; 435 Visitacion Avenue, Brisbane, CA 94005 (US). **HEVEZI, Peter**; 1360 11th Avenue, San Francisco, CA 94122 (US).

(74) Agents: **BASTIAN, Kevin, L.** et al.; Townsend and Townsend and Crew LLP, Two Embarcadero Center, 8th Floor, San Francisco, CA 94111-3834 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/30268 A2

(54) Title: METHODS OF DIAGNOSIS OF PROSTATE CANCER, COMPOSITIONS AND METHODS OF SCREENING FOR MODULATORS OF PROSTATE CANCER

(57) Abstract: Described herein are genes whose expression are up-regulated or down-regulated in prostate cancer. Also described are such genes whose expression is further up-regulated or down-regulated in drug-resistant prostate cancer cells. Related methods and compositions that can be used for diagnosis and treatment of prostate cancer are disclosed. Also described herein are methods that can be used to identify modulators of prostate cancer.

METHODS OF DIAGNOSIS OF PROSTATE CANCER, COMPOSITIONS AND METHODS OF SCREENING FOR MODULATORS OF PROSTATE CANCER

CROSS-REFERENCES TO RELATED APPLICATIONS

This application claims priority from the following applications: USSN 09/687,576 filed October 13, 2000, USSN 60/276,791 filed March 16, 2001; USSN 60/288,589, filed May 4, 2001; USSN 09/733,742, filed December 8, 2000; USSN 09/733,288, filed December 8, 2000; USSN 09/847,046, filed April 30, 2001; USSN 60/276,888, filed March 16, 2001; USSN 60/286,214, filed April 24, 2001; USSN 60/281,922, filed April 6, 2001; USSN 60/263,957, filed January 24, 2001, which are incorporated herein by reference in their entirety.

FIELD OF THE INVENTION

The invention relates to the identification of nucleic acid and protein expression profiles and nucleic acids, products, and antibodies thereto that are involved in prostate cancer; and to the use of such expression profiles and compositions in the diagnosis, prognosis and therapy of prostate cancer. The invention further relates to methods for identifying and using agents and/or targets that inhibit prostate cancer.

BACKGROUND OF THE INVENTION

Prostate cancer is the most commonly diagnosed internal malignancy and second most common cause of cancer death in men in the U.S., resulting in approximately 40,000 deaths each year (Landis et al., *CA Cancer J. Clin.* 48:6-29 (1998); Greenlee et al., *CA Cancer J. Clin.* 50(1):7-13 (2000)), and incidence of prostate cancer has been increasing rapidly over the past 20 years in many parts of the world (Nakata et al., *Int. J. Urol.* 7(7):254-257 (2000); Majeed et al., *BJU Int.* 85(9):1058-1062 (2000)). It develops as the

result of a pathologic transformation of normal prostate cells. In tumorigenesis, the cancer cell undergoes initiation, proliferation and loss of contact inhibition, culminating in invasion of surrounding tissue and, ultimately, metastasis.

Deaths from prostate cancer are a result of metastasis of a prostate tumor. Therefore, early detection of the development of prostate cancer is critical in reducing mortality from this disease. Measuring levels of prostate-specific antigen (PSA) has become a very common method for early detection and screening, and may have contributed to the slight decrease in the mortality rate from prostate cancer in recent years (Nowroozi et al., *Cancer Control* 5(6):522-531 (1998)). However, many cases are not diagnosed until the disease has progressed to an advanced stage.

Treatments such as surgery (prostatectomy) , radiation therapy, and cryotherapy are potentially curative when the cancer remains localized to the prostate. Therefore, early detection of prostate cancer is important for a positive prognosis for treatment. Systemic treatment for metastatic prostate cancer is limited to hormone therapy and chemotherapy. Chemical or surgical castration has been the primary treatment for symptomatic metastatic prostate cancer for over 50 years. This testicular androgen deprivation therapy usually results in stabilization or regression of the disease (in 80% of patients), but progression of metastatic prostate cancer eventually develops (Panvichian et al., *Cancer Control* 3(6):493-500 (1996)). Metastatic disease is currently considered incurable, and the primary goals of treatment are to prolong survival and improve quality of life (Rago, *Cancer Control* 5(6):513-521 (1998)).

Thus, methods that can be used for diagnosis and prognosis of prostate cancer and effective treatment of prostate cancer, and including particularly metastatic prostate cancer, would be desirable. Accordingly, provided herein are methods that can be used in diagnosis and prognosis of prostate cancer. Further provided are methods that can be used to screen candidate bioactive agents for the ability to modulate, e.g., treat, prostate cancer. Additionally, provided herein are molecular targets and compositions for therapeutic intervention in prostate cancer and other cancers.

SUMMARY OF THE INVENTION

The present invention therefore provides nucleotide sequences of genes that are up- and down-regulated in prostate cancer cells. Such genes are useful for diagnostic purposes, and also as targets for screening for therapeutic compounds that modulate prostate cancer, such as hormones or antibodies. Other aspects of the invention will become apparent to the skilled artisan by the following description of the invention.

In one aspect, the present invention provides a method of detecting a prostate cancer-associated transcript in a cell from a patient, the method comprising contacting a biological sample from the patient with a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1-16.

In one embodiment, the present invention provides a method of determining the level of a prostate cancer associated transcript in a cell from a patient.

In one embodiment, the present invention provides a method of detecting a prostate cancer-associated transcript in a cell from a patient, the method comprising contacting a biological sample from the patient with a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1-16.

In one embodiment, the polynucleotide selectively hybridizes to a sequence at least 95% identical to a sequence as shown in Tables 1-16. In another embodiment, the polynucleotide comprises a sequence as shown in Tables 1-16.

In one embodiment, the biological sample is a tissue sample. In another embodiment, the biological sample comprises isolated nucleic acids, e.g., mRNA.

In one embodiment, the polynucleotide is labeled, e.g., with a fluorescent label.

In one embodiment, the polynucleotide is immobilized on a solid surface.

In one embodiment, the patient is undergoing a therapeutic regimen to treat prostate cancer. In another embodiment, the patient is suspected of having metastatic prostate cancer.

In one embodiment, the patient is a human.

In one embodiment, the patient is suspected of having a taxol-resistant cancer.

In one embodiment, the prostate cancer associated transcript is mRNA.

In one embodiment, the method further comprises the step of amplifying nucleic acids before the step of contacting the biological sample with the polynucleotide.

In another aspect, the present invention provides a method of monitoring the efficacy of a therapeutic treatment of prostate cancer, the method comprising the steps of: (i) providing a biological sample from a patient undergoing the therapeutic treatment; and (ii) determining the level of a prostate cancer-associated transcript in the biological sample by contacting the biological sample with a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1-16, thereby monitoring the efficacy of the therapy. In a further embodiment, the patient has metastatic prostate cancer. In a further embodiment, the patient has a drug resistant (e.g., taxol resistant) form of prostate cancer.

In one embodiment, the method further comprises the step of: (iii) comparing the level of the prostate cancer-associated transcript to a level of the prostate cancer-associated transcript in a biological sample from the patient prior to, or earlier in, the therapeutic treatment.

Additionally, provided herein is a method of evaluating the effect of a candidate prostate cancer drug comprising administering the drug to a patient and removing a cell sample from the patient. The expression profile of the cell is then determined. This method may further comprise comparing the expression profile to an expression profile of a healthy individual. In a preferred embodiment, said expression profile includes a gene of Tables 1-16.

In one aspect, the present invention provides an isolated nucleic acid molecule consisting of a polynucleotide sequence as shown in Tables 1-16.

In one embodiment, an expression vector or cell comprises the isolated nucleic acid.

In one aspect, the present invention provides an isolated polypeptide which is encoded by a nucleic acid molecule having polynucleotide sequence as shown in Tables 1-16.

In another aspect, the present invention provides an antibody that specifically binds to an isolated polypeptide which is encoded by a nucleic acid molecule having polynucleotide sequence as shown in Tables 1-16.

In one embodiment, the antibody is conjugated to an effector component, e.g., a fluorescent label, a radioisotope or a cytotoxic chemical.

In one embodiment, the antibody is an antibody fragment. In another embodiment, the antibody is humanized.

In one aspect, the present invention provides a method of detecting a prostate cancer cell in a biological sample from a patient, the method comprising contacting the biological sample with an antibody as described herein.

In another aspect, the present invention provides a method of detecting antibodies specific to prostate cancer in a patient, the method comprising contacting a biological sample from the patient with a polypeptide encoded by a nucleic acid comprising a sequence from Tables 1-16.

In another aspect, the present invention provides a method for identifying a compound that modulates a prostate cancer-associated polypeptide, the method comprising the steps of: (i) contacting the compound with a prostate cancer-associated polypeptide, the polypeptide encoded by a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1-16; and (ii) determining the functional effect of the compound upon the polypeptide.

In one embodiment, the functional effect is a physical effect, an enzymatic effect, or a chemical effect.

In one embodiment, the polypeptide is expressed in a eukaryotic host cell or cell membrane. In another embodiment, the polypeptide is recombinant.

In one embodiment, the functional effect is determined by measuring ligand binding to the polypeptide.

In another aspect, the present invention provides a method of inhibiting proliferation of a prostate cancer-associated cell to treat prostate cancer in a patient, the method comprising the step of administering to the subject a therapeutically effective amount of a compound identified as described herein.

In one embodiment, the compound is an antibody.

In another aspect, the present invention provides a drug screening assay comprising the steps of: (i) administering a test compound to a mammal having prostate cancer or to a cell sample isolated therefrom; (ii) comparing the level of gene expression of a

polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1-16 in a treated cell or mammal with the level of gene expression of the polynucleotide in a control cell sample or mammal, wherein a test compound that modulates the level of expression of the polynucleotide is a candidate for the treatment of prostate cancer.

In one embodiment, the control is a mammal with prostate cancer or a cell sample therefrom that has not been treated with the test compound. In another embodiment, the control is a normal cell or mammal.

In one embodiment, the test compound is administered in varying amounts or concentrations. In another embodiment, the test compound is administered for varying time periods. In another embodiment, the comparison can occur after addition or removal of the drug candidate.

In one embodiment, the levels of a plurality of polynucleotides that selectively hybridize to a sequence at least 80% identical to a sequence as shown in Tables 1-16 are individually compared to their respective levels in a control cell sample or mammal. In a preferred embodiment the plurality of polynucleotides is from three to ten.

In another aspect, the present invention provides a method for treating a mammal having prostate cancer comprising administering a compound identified by the assay described herein.

In another aspect, the present invention provides a pharmaceutical composition for treating a mammal having prostate cancer, the composition comprising a compound identified by the assay described herein and a physiologically acceptable excipient.

In one aspect, the present invention provides a method of screening drug candidates by providing a cell expressing a gene that is up- and down-regulated as in a prostate cancer. In one embodiment, a gene is selected from Tables 1-16. The method further includes adding a drug candidate to the cell and determining the effect of the drug candidate on the expression of the expression profile gene.

In one embodiment, the method of screening drug candidates includes comparing the level of expression in the absence of the drug candidate to the level of expression in the presence of the drug candidate, wherein the concentration of the drug

candidate can vary when present, and wherein the comparison can occur after addition or removal of the drug candidate. In a preferred embodiment, the cell expresses at least two expression profile genes. The profile genes may show an increase or decrease.

Also provided is a method of evaluating the effect of a candidate prostate cancer drug comprising administering the drug to a transgenic animal expressing or over-expressing the prostate cancer modulatory protein, or an animal lacking the prostate cancer modulatory protein, for example as a result of a gene knockout.

Moreover, provided herein is a biochip comprising one or more nucleic acid segments of Tables 1-16, wherein the biochip comprises fewer than 1000 nucleic acid probes. Preferably, at least two nucleic acid segments are included. More preferably, at least three nucleic acid segments are included.

Furthermore, a method of diagnosing a disorder associated with prostate cancer is provided. The method comprises determining the expression of a gene of Tables 1-16, in a first tissue type of a first individual, and comparing the distribution to the expression of the gene from a second normal tissue type from the first individual or a second unaffected individual. A difference in the expression indicates that the first individual has a disorder associated with prostate cancer.

In a further embodiment, the biochip also includes a polynucleotide sequence of a gene that is not up- and down-regulated in prostate cancer.

In one embodiment a method for screening for a bioactive agent capable of interfering with the binding of a prostate cancer modulating protein (prostate cancer modulatory protein) or a fragment thereof and an antibody which binds to said prostate cancer modulatory protein or fragment thereof. In a preferred embodiment, the method comprises combining a prostate cancer modulatory protein or fragment thereof, a candidate bioactive agent and an antibody which binds to said prostate cancer modulatory protein or fragment thereof. The method further includes determining the binding of said prostate cancer modulatory protein or fragment thereof and said antibody. Wherein there is a change in binding, an agent is identified as an interfering agent. The interfering agent can be an agonist or an antagonist. Preferably, the agent inhibits prostate cancer.

Also provided herein are methods of eliciting an immune response in an individual. In one embodiment a method provided herein comprises administering to an

individual a composition comprising a prostate cancer modulating protein, or a fragment thereof. In another embodiment, the protein is encoded by a nucleic acid selected from those of Tables 1-16.

Further provided herein are compositions capable of eliciting an immune response in an individual. In one embodiment, a composition provided herein comprises a prostate cancer modulating protein, preferably encoded by a nucleic acid of Tables 1-16, or a fragment thereof, and a pharmaceutically acceptable carrier. In another embodiment, said composition comprises a nucleic acid comprising a sequence encoding a prostate cancer modulating protein, preferably selected from the nucleic acids of Tables 1-16, and a pharmaceutically acceptable carrier.

Also provided are methods of neutralizing the effect of a prostate cancer protein, or a fragment thereof, comprising contacting an agent specific for said protein with said protein in an amount sufficient to effect neutralization. In another embodiment, the protein is encoded by a nucleic acid selected from those of Tables 1-16.

In another aspect of the invention, a method of treating an individual for prostate cancer is provided. In one embodiment, the method comprises administering to said individual an inhibitor of a prostate cancer modulating protein. In another embodiment, the method comprises administering to a patient having prostate cancer an antibody to a prostate cancer modulating protein conjugated to a therapeutic moiety. Such a therapeutic moiety can be a cytotoxic agent or a radioisotope.

DETAILED DESCRIPTION OF THE INVENTION

In accordance with the objects outlined above, the present invention provides novel methods for diagnosis and prognosis evaluation for prostate cancer (PC), including metastatic prostate cancer, as well as methods for screening for compositions which modulate prostate cancer. Also provided are methods for treating prostate cancer.

In addition to the other nucleic acid and peptide sequences, the present invention also relates to the identification of PAA2 as a gene that is highly over expressed in prostate cancer patient tissues. PAA2 sequence is identical to the zinc transporter ZNT4. Results presented herein demonstrate that PAA2/ZNT4 is highly expressed in prostate cancer cells. The prostate gland is unique in that it has the highest capacity of any organ in the body

to accumulate zinc. Zinc uptake is regulated by prolactin and testosterone, which induce the expression of a member of the ZIP family of zinc transporters (Costello et al., 1999, J. Biol. Chem. 274:17499-17504). Zinc accumulation in the prostate functions to inhibit citrate oxidation, which results in a decrease in cellular ATP production (Costello and Franklin, 1998, Prostate 35:285-296). Cancer cells are more sensitive to decreased ATP production and have evolved to prevent zinc accumulation. Without wishing to be bound by theory, the up-regulation of ZNT4 in prostate cancer cells may result in protection of the cells from high zinc levels by its ability to pump accumulated zinc out of the cells.

The present invention also relates to nucleic acid sequencess encoding PBH1. PBH1 is related to human TRPC7 (transient receptor potential-related channels, NP_003298), a putative calcium channel highly expressed in brain (Nagamine et al., Genomics 54:124-131 (1998)). Trp is related to melastatin, a gene down-regulated in metastatic melanomas (Duncan et al., Cancer Res. 58:1515-1520 (1998)), and MTR1, a gene locallized to within the Beckwith-Wiedemann syndrome/Wilm's tumor susceptability region (Prawitt et al., Hum. Mol. Genet. 9:203-216 (2000)). Without wishing to be bound by theory, it is believed that PBH1 functions as a calcium channel.

As a calcium channel, PBH1 is an ideal target for a small molecule therapeutic, or a therapeutic antibody that disrupts channel function. CD20, the target of Rituximab in non-Hodgekin's lymphoma (Maloney et al., Blood 90:2188-2195 (1997); Leget and Czuczman, Curr. Opin. Oncol. 10:548-551 (1998)), is a plasma membrane calcium channel expressed in B cells (Tedder and Engel, Immunol. Today 15:450-454 (1994)). Similarly, a small molecule, or antibody that inhibits or alters a calcium signal mediated by PBH1, will result in the death of prostate cancer cells.

PBH1, and other genes of the invention, are also be useful as targets for cytotoxic T-lymphocytes. Genes that are tumor specific, or that are expressed in immune-privileged organs, are currently being used as potential vaccine targets (Van den Eynde and Boon, Int. J. Clin. Lab. Res. 27:81-86 (1997)). The expression pattern of PBH1 indicates that it is an ideal target for cytotoxic T-lymphocytes. Thus, therapies that utilize PBH1-specific cytotoxic T-lymphocytes to induce prostate cancer cell death are also provided by this invention. See, e.g., U.S. Patent No. 6,051,227 and WO 00/32231, the disclosures of which are herein incorporated by reference.

The present invention is also related to the identification of PAA3 as a gene that is important in the modulation of prostate cancer and or breast cancer.

Tables 1-16 provide unigene cluster identification numbers, exemplar accession numbers, or genomic nucleotide position numbers for the nucleotide sequence of genes that exhibit increased or decreased expression in prostate cancer samples.

Definitions

The term "prostate cancer protein" or "prostate cancer polynucleotide" or "prostate cancer-associated transcript" refers to nucleic acid and polypeptide polymorphic variants, alleles, mutants, and interspecies homologues that: (1) have a nucleotide sequence that has greater than about 60% nucleotide sequence identity, 65%, 70%, 75%, 80%, 85%, 90%, preferably 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% or greater nucleotide sequence identity, preferably over a region of over a region of at least about 25, 50, 100, 200, 500, 1000, or more nucleotides, to a nucleotide sequence of or associated with a unigene cluster of Tables 1-16; (2) bind to antibodies, e.g., polyclonal antibodies, raised against an immunogen comprising an amino acid sequence encoded by a nucleotide sequence of or associated with a unigene cluster of Tables 1-16, and conservatively modified variants thereof; (3) specifically hybridize under stringent hybridization conditions to a nucleic acid sequence, or the complement thereof of Tables 1-16 and conservatively modified variants thereof or (4) have an amino acid sequence that has greater than about 60% amino acid sequence identity, 65%, 70%, 75%, 80%, 85%, 90%, preferably 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% or greater amino sequence identity, preferably over a region of over a region of at least about 25, 50, 100, 200, 500, 1000, or more amino acid, to an amino acid sequence encoded by a nucleotide sequence of or associated with a unigene cluster of Tables 1-16. A polynucleotide or polypeptide sequence is typically from a mammal including, but not limited to, primate, e.g., human; rodent, e.g., rat, mouse, hamster; cow, pig, horse, sheep, or other mammal. A "prostate cancer polypeptide" and a "prostate cancer polynucleotide," include both naturally occurring or recombinant forms.

A "full length" prostate cancer protein or nucleic acid refers to a prostate cancer polypeptide or polynucleotide sequence, or a variant thereof, that contains all of the elements normally contained in one or more naturally occurring, wild type prostate cancer

polynucleotide or polypeptide sequences. For example, a full length prostate cancer nucleic acid will typically comprise all of the exons that encode for the full length, naturally ocurring protein. The "full length" may be prior to, or after, various stages of post-translation processing or splicing, including alternative splicing.

"Biological sample" as used herein is a sample of biological tissue or fluid that contains nucleic acids or polypeptides, e.g., of a prostate cancer protein, polynucleotide or transcript. Such samples include, but are not limited to, tissue isolated from primates, e.g., humans, or rodents, e.g., mice, and rats. Biological samples may also include sections of tissues such as biopsy and autopsy samples, frozen sections taken for histologic purposes, blood, plasma, serum, sputum, stool, tears, mucus, hair, skin, etc. Biological samples also include explants and primary and/or transformed cell cultures derived from patient tissues. A biological sample is typically obtained from a eukaryotic organism, most preferably a mammal such as a primate e.g., chimpanzee or human; cow; dog; cat; a rodent, e.g., guinea pig, rat, mouse; rabbit; or a bird; reptile; or fish.

"Providing a biological sample" means to obtain a biological sample for use in methods described in this invention. Most often, this will be done by removing a sample of cells from an animal, but can also be accomplished by using previously isolated cells (*e.g.*, isolated by another person, at another time, and/or for another purpose), or by performing the methods of the invention *in vivo*. Archival tissues, having treatment or outcome history, will be particularly useful.

The terms "identical" or percent "identity," in the context of two or more nucleic acids or polypeptide sequences, refer to two or more sequences or subsequences that are the same or have a specified percentage of amino acid residues or nucleotides that are the same (i.e., about 60% identity, preferably 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or higher identity over a specified region, when compared and aligned for maximum correspondence over a comparison window or designated region) as measured using a BLAST or BLAST 2.0 sequence comparison algorithms with default parameters described below, or by manual alignment and visual inspection (*see, e.g.*, NCBI web site http://www.ncbi.nlm.nih.gov/BLAST/ or the like). Such sequences are then said to be "substantially identical." This definition also refers to, or may be applied to, the compliment of a test sequence. The definition also includes sequences that have deletions

and/or additions, as well as those that have substitutions, as well as naturally occurring, e.g., polymorphic or allelic variants, and man-made variants. As described below, the preferred algorithms can account for gaps and the like. Preferably, identity exists over a region that is at least about 25 amino acids or nucleotides in length, or more preferably over a region that is 50-100 amino acids or nucleotides in length.

For sequence comparison, typically one sequence acts as a reference sequence, to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are entered into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. Preferably, default program parameters can be used, or alternative parameters can be designated. The sequence comparison algorithm then calculates the percent sequence identities for the test sequences relative to the reference sequence, based on the program parameters.

A "comparison window", as used herein, includes reference to a segment of one of the number of contiguous positions selected from the group consisting typically of from 20 to 600, usually about 50 to about 200, more usually about 100 to about 150 in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned. Methods of alignment of sequences for comparison are well-known in the art. Optimal alignment of sequences for comparison can be conducted, e.g., by the local homology algorithm of Smith & Waterman, *Adv. Appl. Math.* 2:482 (1981), by the homology alignment algorithm of Needleman & Wunsch, *J. Mol. Biol.* 48:443 (1970), by the search for similarity method of Pearson & Lipman, *Proc. Nat'l. Acad. Sci. USA* 85:2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), or by manual alignment and visual inspection (*see, e.g., Current Protocols in Molecular Biology* (Ausubel *et al.*, eds. 1995 supplement)).

Preferred examples of algorithms that are suitable for determining percent sequence identity and sequence similarity include the BLAST and BLAST 2.0 algorithms, which are described in Altschul *et al., Nuc. Acids Res.* 25:3389-3402 (1977) and Altschul *et al., J. Mol. Biol.* 215:403-410 (1990). BLAST and BLAST 2.0 are used, with the parameters described herein, to determine percent sequence identity for the nucleic acids and proteins of

the invention. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/). This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul *et al.*, *supra*). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, e.g., for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues; always < 0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, M=5, N=-4 and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength of 3, and expectation (E) of 10, and the BLOSUM62 scoring matrix (*see* Henikoff & Henikoff, *Proc. Natl. Acad. Sci. USA* 89:10915 (1989)) alignments (B) of 50, expectation (E) of 10, M=5, N=-4, and a comparison of both strands.

The BLAST algorithm also performs a statistical analysis of the similarity between two sequences (*see, e.g.,* Karlin & Altschul, *Proc. Nat'l. Acad. Sci. USA* 90:5873-5787 (1993)). One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a reference sequence if the smallest sum probability in a comparison of the test nucleic acid to the reference nucleic acid is less than about 0.2, more preferably less than about 0.01, and most preferably less than about 0.001. Log values may be large negative numbers, e.g., 5, 10, 20, 30, 40, 40, 70, 90, 110, 150, 170, etc.

An indication that two nucleic acid sequences or polypeptides are substantially identical is that the polypeptide encoded by the first nucleic acid is immunologically cross reactive with the antibodies raised against the polypeptide encoded by the second nucleic acid, as described below. Thus, a polypeptide is typically substantially identical to a second polypeptide, e.g., where the two peptides differ only by conservative substitutions. Another indication that two nucleic acid sequences are substantially identical is that the two molecules or their complements hybridize to each other under stringent conditions, as described below. Yet another indication that two nucleic acid sequences are substantially identical is that the same primers can be used to amplify the sequences.

A "host cell" is a naturally occurring cell or a transformed cell that contains an expression vector and supports the replication or expression of the expression vector. Host cells may be cultured cells, explants, cells *in vivo*, and the like. Host cells may be prokaryotic cells such as *E. coli*, or eukaryotic cells such as yeast, insect, amphibian, or mammalian cells such as CHO, HeLa, and the like (*see, e.g.,* the American Type Culture Collection catalog or web site, www.atcc.org).

The terms "isolated," "purified," or "biologically pure" refer to material that is substantially or essentially free from components that normally accompany it as found in its native state. Purity and homogeneity are typically determined using analytical chemistry techniques such as polyacrylamide gel electrophoresis or high performance liquid chromatography. A protein or nucleic acid that is the predominant species present in a preparation is substantially purified. In particular, an isolated nucleic acid is separated from some open reading frames that naturally flank the gene and encode proteins other than protein encoded by the gene. The term "purified" in some embodiments denotes that a nucleic acid or protein gives rise to essentially one band in an electrophoretic gel. Preferably, it means that the nucleic acid or protein is at least 85% pure, more preferably at least 95% pure, and most preferably at least 99% pure. "Purify" or "purification" in other embodiments means removing at least one contaminant from the composition to be purified. In this sense, purification does not require that the purified compound be homogenous, e.g., 100% pure.

The terms "polypeptide," "peptide" and "protein" are used interchangeably herein to refer to a polymer of amino acid residues. The terms apply to amino acid polymers in which one or more amino acid residue is an artificial chemical mimetic of a corresponding

naturally occurring amino acid, as well as to naturally occurring amino acid polymers, those containing modified residues, and non-naturally occurring amino acid polymer.

The term "amino acid" refers to naturally occurring and synthetic amino acids, as well as amino acid analogs and amino acid mimetics that function similarly to the naturally occurring amino acids. Naturally occurring amino acids are those encoded by the genetic code, as well as those amino acids that are later modified, e.g., hydroxyproline, γ-carboxyglutamate, and O-phosphoserine. Amino acid analogs refers to compounds that have the same basic chemical structure as a naturally occurring amino acid, e.g., an α carbon that is bound to a hydrogen, a carboxyl group, an amino group, and an R group, e.g., homoserine, norleucine, methionine sulfoxide, methionine methyl sulfonium. Such analogs may have modified R groups (e.g., norleucine) or modified peptide backbones, but retain the same basic chemical structure as a naturally occurring amino acid. Amino acid mimetics refers to chemical compounds that have a structure that is different from the general chemical structure of an amino acid, but that functions similarly to a naturally occurring amino acid.

Amino acids may be referred to herein by either their commonly known three letter symbols or by the one-letter symbols recommended by the IUPAC-IUB Biochemical Nomenclature Commission. Nucleotides, likewise, may be referred to by their commonly accepted single-letter codes.

"Conservatively modified variants" applies to both amino acid and nucleic acid sequences. With respect to particular nucleic acid sequences, conservatively modified variants refers to those nucleic acids which encode identical or essentially identical amino acid sequences, or where the nucleic acid does not encode an amino acid sequence, to essentially identical or associated, e.g., naturally contiguous, sequences. Because of the degeneracy of the genetic code, a large number of functionally identical nucleic acids encode most proteins. For instance, the codons GCA, GCC, GCG and GCU all encode the amino acid alanine. Thus, at every position where an alanine is specified by a codon, the codon can be altered to another of the corresponding codons described without altering the encoded polypeptide. Such nucleic acid variations are "silent variations," which are one species of conservatively modified variations. Every nucleic acid sequence herein which encodes a polypeptide also describes silent variations of the nucleic acid. One of skill will recognize that in certain contexts each codon in a nucleic acid (except AUG, which is ordinarily the

only codon for methionine, and TGG, which is ordinarily the only codon for tryptophan) can be modified to yield a functionally identical molecule. Accordingly, often silent variations of a nucleic acid which encodes a polypeptide is implicit in a described sequence with respect to the expression product, but not with respect to actual probe sequences.

As to amino acid sequences, one of skill will recognize that individual substitutions, deletions or additions to a nucleic acid, peptide, polypeptide, or protein sequence which alters, adds or deletes a single amino acid or a small percentage of amino acids in the encoded sequence is a "conservatively modified variant" where the alteration results in the substitution of an amino acid with a chemically similar amino acid. Conservative substitution tables providing functionally similar amino acids are well known in the art. Such conservatively modified variants are in addition to and do not exclude polymorphic variants, interspecies homologs, and alleles of the invention.typically conservative substitutions for one another: 1) Alanine (A), Glycine (G); 2) Aspartic acid (D), Glutamic acid (E); 3) Asparagine (N), Glutamine (Q); 4) Arginine (R), Lysine (K); 5) Isoleucine (I), Leucine (L), Methionine (M), Valine (V); 6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W); 7) Serine (S), Threonine (T); and 8) Cysteine (C), Methionine (M) (*see, e.g.*, Creighton, *Proteins* (1984)).

Macromolecular structures such as polypeptide structures can be described in terms of various levels of organization. For a general discussion of this organization, *see, e.g.*, Alberts *et al.*, *Molecular Biology of the Cell* (3rd ed., 1994) and Cantor & Schimmel, *Biophysical Chemistry Part I: The Conformation of Biological Macromolecules* (1980). "Primary structure" refers to the amino acid sequence of a particular peptide. "Secondary structure" refers to locally ordered, three dimensional structures within a polypeptide. These structures are commonly known as domains. Domains are portions of a polypeptide that often form a compact unit of the polypeptide and are typically 25 to approximately 500 amino acids long. Typical domains are made up of sections of lesser organization such as stretches of β-sheet and α-helices. "Tertiary structure" refers to the complete three dimensional structure of a polypeptide monomer. "Quaternary structure" refers to the three dimensional structure formed, usually by the noncovalent association of independent tertiary units. Anisotropic terms are also known as energy terms.

"Nucleic acid" or "oligonucleotide" or "polynucleotide" or grammatical equivalents used herein means at least two nucleotides covalently linked together. Oligonucleotides are typically from about 5, 6, 7, 8, 9, 10, 12, 15, 25, 30, 40, 50 or more nucleotides in length, up to about 100 nucleotides in length. Nucleic acids and polynucleotides are a polymers of any length, including longer lengths, e.g., 200, 300, 500, 1000, 2000, 3000, 5000, 7000, 10,000, etc. A nucleic acid of the present invention will generally contain phosphodiester bonds, although in some cases, nucleic acid analogs are included that may have alternate backbones, comprising, e.g., phosphoramidate, phosphorothioate, phosphorodithioate, or O-methylphophoroamidite linkages (see Eckstein, Oligonucleotides and Analogues: A Practical Approach, Oxford University Press); and peptide nucleic acid backbones and linkages. Other analog nucleic acids include those with positive backbones; non-ionic backbones, and non-ribose backbones, including those described in U.S. Patent Nos. 5,235,033 and 5,034,506, and Chapters 6 and 7, ASC Symposium Series 580, *Carbohydrate Modifications in Antisense Research*, Sanghui & Cook, eds.. Nucleic acids containing one or more carbocyclic sugars are also included within one definition of nucleic acids. Modifications of the ribose-phosphate backbone may be done for a variety of reasons, e.g. to increase the stability and half-life of such molecules in physiological environments or as probes on a biochip. Mixtures of naturally occurring nucleic acids and analogs can be made; alternatively, mixtures of different nucleic acid analogs, and mixtures of naturally occurring nucleic acids and analogs may be made.

A variety of references disclose such nucleic acid analogs, including, for example, phosphoramidate (Beaucage et al., Tetrahedron 49(10):1925 (1993) and references therein; Letsinger, J. Org. Chem. 35:3800 (1970); Sprinzl et al., Eur. J. Biochem. 81:579 (1977); Letsinger et al., Nucl. Acids Res. 14:3487 (1986); Sawai et al, Chem. Lett. 805 (1984), Letsinger et al., J. Am. Chem. Soc. 110:4470 (1988); and Pauwels et al., Chemica Scripta 26:141 91986)), phosphorothioate (Mag et al., Nucleic Acids Res. 19:1437 (1991); and U.S. Patent No. 5,644,048), phosphorodithioate (Briu et al., J. Am. Chem. Soc. 111:2321 (1989), O-methylphophoroamidite linkages (see Eckstein, Oligonucleotides and Analogues: A Practical Approach, Oxford University Press), and peptide nucleic acid backbones and linkages (see Egholm, J. Am. Chem. Soc. 114:1895 (1992); Meier et al., Chem. Int. Ed. Engl. 31:1008 (1992); Nielsen, Nature, 365:566 (1993); Carlsson et al., Nature 380:207 (1996), all

of which are incorporated by reference). Other analog nucleic acids include those with positive backbones (Denpcy et al., Proc. Natl. Acad. Sci. USA 92:6097 (1995); non-ionic backbones (U.S. Patent Nos. 5,386,023, 5,637,684, 5,602,240, 5,216,141 and 4,469,863; Kiedrowshi et al., Angew. Chem. Intl. Ed. English 30:423 (1991); Letsinger et al., J. Am. Chem. Soc. 110:4470 (1988); Letsinger et al., Nucleoside & Nucleotide 13:1597 (1994); Chapters 2 and 3, ASC Symposium Series 580, "Carbohydrate Modifications in Antisense Research", Ed. Y.S. Sanghui and P. Dan Cook; Mesmaeker et al., Bioorganic & Medicinal Chem. Lett. 4:395 (1994); Jeffs et al., J. Biomolecular NMR 34:17 (1994); Tetrahedron Lett. 37:743 (1996)) and non-ribose backbones, including those described in U.S. Patent Nos. 5,235,033 and 5,034,506, and Chapters 6 and 7, ASC Symposium Series 580, "Carbohydrate Modifications in Antisense Research", Ed. Y.S. Sanghui and P. Dan Cook. Nucleic acids containing one or more carbocyclic sugars are also included within one definition of nucleic acids (see Jenkins et al., Chem. Soc. Rev. (1995) pp 169-176). Several nucleic acid analogs are described in Rawls, C & E News June 2, 1997 page 35. All of these references are hereby expressly incorporated by reference.

Particularly preferred are peptide nucleic acids (PNA) which includes peptide nucleic acid analogs. These backbones are substantially non-ionic under neutral conditions, in contrast to the highly charged phosphodiester backbone of naturally occurring nucleic acids. This results in two advantages. First, the PNA backbone exhibits improved hybridization kinetics. PNAs have larger changes in the melting temperature (T_m) for mismatched versus perfectly matched basepairs. DNA and RNA typically exhibit a 2-4°C drop in T_m for an internal mismatch. With the non-ionic PNA backbone, the drop is closer to 7-9°C. Similarly, due to their non-ionic nature, hybridization of the bases attached to these backbones is relatively insensitive to salt concentration. In addition, PNAs are not degraded by cellular enzymes, and thus can be more stable.

The nucleic acids may be single stranded or double stranded, as specified, or contain portions of both double stranded or single stranded sequence. As will be appreciated by those in the art, the depiction of a single strand also defines the sequence of the complementary strand; thus the sequences described herein also provide the complement of the sequence. The nucleic acid may be DNA, both genomic and cDNA, RNA or a hybrid, where the nucleic acid may contain combinations of deoxyribo- and ribo-nucleotides, and

combinations of bases, including uracil, adenine, thymine, cytosine, guanine, inosine, xanthine hypoxanthine, isocytosine, isoguanine, etc. "Transcript" typically refers to a naturally occurring RNA, e.g., a pre-mRNA, hnRNA, or mRNA. As used herein, the term "nucleoside" includes nucleotides and nucleoside and nucleotide analogs, and modified nucleosides such as amino modified nucleosides. In addition, "nucleoside" includes non-naturally occurring analog structures. Thus, e.g. the individual units of a peptide nucleic acid, each containing a base, are referred to herein as a nucleoside.

A "label" or a "detectable moiety" is a composition detectable by spectroscopic, photochemical, biochemical, immunochemical, chemical, or other physical means. For example, useful labels include fluorescent dyes, electron-dense reagents, enzymes (*e.g.*, as commonly used in an ELISA), biotin, digoxigenin, or haptens and proteins or other entities which can be made detectable, *e.g.*, by incorporating a radiolabel into the peptide or used to detect antibodies specifically reactive with the peptide. The radioisotope may be, for example, 3H, 14C, 32P, 35S, or 125I. In some cases, particularly using antibodies against the proteins of the invention, the radioisotopes are used as toxic moieties, as described below. The labels may be incorporated into the prostate cancer nucleic acids, proteins and antibodies at any position. Any method known in the art for conjugating the antibody to the label may be employed, including those methods described by Hunter et al., Nature, 144:945 (1962); David et al., Biochemistry, 13:1014 (1974); Pain et al., J. Immunol. Meth., 40:219 (1981); and Nygren, J. Histochem. and Cytochem., 30:407 (1982). The lifetime of radiolabeled peptides or radiolabeled antibody compositions may extended by the addition of substances that stablize the radiolabeled peptide or antibody and protect it from degradation. Any substance or combination of substances that stablize the radiolabeled peptide or antibody may be used including those substances disclosed in US Patent No. 5,961,955.

An "effector" or "effector moiety" or "effector component" is a molecule that is bound (or linked, or conjugated), either covalently, through a linker or a chemical bond, or noncovalently, through ionic, van der Waals, electrostatic, or hydrogen bonds, to an antibody. The "effector" can be a variety of molecules including, e.g., detection moieties including radioactive compounds, fluorescent compounds, an enzyme or substrate, tags such as epitope tags, a toxin; activatable moieties, a chemotherapeutic agent; a lipase; an antibiotic; or a radioisotope emitting "hard" e.g., beta radiation.

A "labeled nucleic acid probe or oligonucleotide" is one that is bound, either covalently, through a linker or a chemical bond, or noncovalently, through ionic, van der Waals, electrostatic, or hydrogen bonds to a label such that the presence of the probe may be detected by detecting the presence of the label bound to the probe. Alternatively, method using high affinity interactions may achieve the same results where one of a pair of binding partners binds to the other, e.g., biotin, streptavidin.

As used herein a "nucleic acid probe or oligonucleotide" is defined as a nucleic acid capable of binding to a target nucleic acid of complementary sequence through one or more types of chemical bonds, usually through complementary base pairing, usually through hydrogen bond formation. As used herein, a probe may include natural (i.e., A, G, C, or T) or modified bases (7-deazaguanosine, inosine, etc.). In addition, the bases in a probe may be joined by a linkage other than a phosphodiester bond, so long as it does not functionally interfere with hybridization. Thus, e.g., probes may be peptide nucleic acids in which the constituent bases are joined by peptide bonds rather than phosphodiester linkages. It will be understood by one of skill in the art that probes may bind target sequences lacking complete complementarity with the probe sequence depending upon the stringency of the hybridization conditions. The probes are preferably directly labeled as with isotopes, chromophores, lumiphores, chromogens, or indirectly labeled such as with biotin to which a streptavidin complex may later bind. By assaying for the presence or absence of the probe, one can detect the presence or absence of the select sequence or subsequence. Diagnosis or prognosis may be based at the genomic level, or at the level of RNA or protein expression.

The term "recombinant" when used with reference, e.g., to a cell, or nucleic acid, protein, or vector, indicates that the cell, nucleic acid, protein or vector, has been modified by the introduction of a heterologous nucleic acid or protein or the alteration of a native nucleic acid or protein, or that the cell is derived from a cell so modified. Thus, e.g., recombinant cells express genes that are not found within the native (non-recombinant) form of the cell or express native genes that are otherwise abnormally expressed, under expressed or not expressed at all. By the term "recombinant nucleic acid" herein is meant nucleic acid, originally formed *in vitro*, in general, by the manipulation of nucleic acid, e.g., using polymerases and endonucleases, in a form not normally found in nature. In this manner, operably linkage of different sequences is achieved. Thus an isolated nucleic acid, in a linear

form, or an expression vector formed *in vitro* by ligating DNA molecules that are not normally joined, are both considered recombinant for the purposes of this invention. It is understood that once a recombinant nucleic acid is made and reintroduced into a host cell or organism, it will replicate non-recombinantly, i.e., using the *in vivo* cellular machinery of the host cell rather than *in vitro* manipulations; however, such nucleic acids, once produced recombinantly, although subsequently replicated non-recombinantly, are still considered recombinant for the purposes of the invention. Similarly, a "recombinant protein" is a protein made using recombinant techniques, i.e., through the expression of a recombinant nucleic acid as depicted above.

The term "heterologous" when used with reference to portions of a nucleic acid indicates that the nucleic acid comprises two or more subsequences that are not normally found in the same relationship to each other in nature. For instance, the nucleic acid is typically recombinantly produced, having two or more sequences, e.g., from unrelated genes arranged to make a new functional nucleic acid, e.g., a promoter from one source and a coding region from another source. Similarly, a heterologous protein will often refer to two or more subsequences that are not found in the same relationship to each other in nature (e.g., a fusion protein).

A "promoter" is defined as an array of nucleic acid control sequences that direct transcription of a nucleic acid. As used herein, a promoter includes necessary nucleic acid sequences near the start site of transcription, such as, in the case of a polymerase II type promoter, a TATA element. A promoter also optionally includes distal enhancer or repressor elements, which can be located as much as several thousand base pairs from the start site of transcription. A "constitutive" promoter is a promoter that is active under most environmental and developmental conditions. An "inducible" promoter is a promoter that is active under environmental or developmental regulation. The term "operably linked" refers to a functional linkage between a nucleic acid expression control sequence (such as a promoter, or array of transcription factor binding sites) and a second nucleic acid sequence, wherein the expression control sequence directs transcription of the nucleic acid corresponding to the second sequence.

An "expression vector" is a nucleic acid construct, generated recombinantly or synthetically, with a series of specified nucleic acid elements that permit transcription of a

particular nucleic acid in a host cell. The expression vector can be part of a plasmid, virus, or nucleic acid fragment. Typically, the expression vector includes a nucleic acid to be transcribed operably linked to a promoter.

The phrase "selectively (or specifically) hybridizes to" refers to the binding, duplexing, or hybridizing of a molecule only to a particular nucleotide sequence that is determinative of the presence of the nucleotide sequence, in a heterogeneous population of nucleic acids and other biologics (e.g., total cellular or library DNA or RNA). Similarly, the phrase "specifically (or selectively) binds" to an antibody or "specifically (or selectively) immunoreactive with," when referring to a protein or peptide, refers to a binding reaction that is determinative of the presence of the protein, in a heterogeneous population of proteins and other biologics. Thus, under designated immunoassay or nucleic acid hybridization conditions, the specified antibodies or nucleic acid probes bind to a particular protein nucleotide sequences at least two times the background and more typically more than 10 to 100 times background.

Specific binding to an antibody under such conditions requires an antibody that is selected for its specificity for a particular protein. For example, polyclonal antibodies raised to a particular protein, polymorphic variants, alleles, orthologs, and conservatively modified variants, or splice variants, or portions thereof, can be selected to obtain only those polyclonal antibodies that are specifically immunoreactive with the desired prostact cancer protein and not with other proteins. This selection may be achieved by subtracting out antibodies that cross-react with other molecules. A variety of immunoassay formats may be used to select antibodies specifically immunoreactive with a particular protein. For example, solid-phase ELISA immunoassays are routinely used to select antibodies specifically immunoreactive with a protein (*see, e.g.*, Harlow & Lane, *Antibodies, A Laboratory Manual* (1988) for a description of immunoassay formats and conditions that can be used to determine specific immunoreactivity).

The phrase "stringent hybridization conditions" refers to conditions under which a probe will hybridize to its target subsequence, typically in a complex mixture of nucleic acids, but to no other sequences. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures. An extensive guide to the hybridization of nucleic acids is found in

Tijssen, *Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Probes*, "Overview of principles of hybridization and the strategy of nucleic acid assays" (1993). Generally, stringent conditions are selected to be about 5-10°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength pH. The T_m is the temperature (under defined ionic strength, pH, and nucleic concentration) at which 50% of the probes complementary to the target hybridize to the target sequence at equilibrium (as the target sequences are present in excess, at T_m, 50% of the probes are occupied at equilibrium). Stringent conditions will be those in which the salt concentration is less than about 1.0 M sodium ion, typically about 0.01 to 1.0 M sodium ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes (e.g., 10 to 50 nucleotides) and at least about 60°C for long probes (e.g., greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. For selective or specific hybridization, a positive signal is at least two times background, preferably 10 times background hybridization. Exemplary stringent hybridization conditions can be as following: 50% formamide, 5x SSC, and 1% SDS, incubating at 42°C, or, 5x SSC, 1% SDS, incubating at 65°C, with wash in 0.2x SSC, and 0.1% SDS at 65°C. For PCR, a temperature of about 36°C is typical for low stringency amplification, although annealing temperatures may vary between about 32°C and 48°C depending on primer length. For high stringency PCR amplification, a temperature of about 62°C is typical, although high stringency annealing temperatures can range from about 50°C to about 65°C, depending on the primer length and specificity. Typical cycle conditions for both high and low stringency amplifications include a denaturation phase of 90°C - 95°C for 30 sec - 2 min., an annealing phase lasting 30 sec. - 2 min., and an extension phase of about 72°C for 1 - 2 min. Protocols and guidelines for low and high stringency amplification reactions are provided, e.g., in Innis *et al.* (1990) *PCR Protocols, A Guide to Methods and Applications*, Academic Press, Inc. N.Y.).

Nucleic acids that do not hybridize to each other under stringent conditions are still substantially identical if the polypeptides which they encode are substantially identical. This occurs, e.g., when a copy of a nucleic acid is created using the maximum codon degeneracy permitted by the genetic code. In such cases, the nucleic acids typically hybridize

under moderately stringent hybridization conditions. Exemplary "moderately stringent hybridization conditions" include a hybridization in a buffer of 40% formamide, 1 M NaCl, 1% SDS at 37°C, and a wash in 1X SSC at 45°C. A positive hybridization is at least twice background. Those of ordinary skill will readily recognize that alternative hybridization and wash conditions can be utilized to provide conditions of similar stringency. Additional guidelines for determining hybridization parameters are provided in numerous reference, e.g., and Current Protocols in Molecular Biology, ed. Ausubel, *et al.*

The phrase "functional effects" in the context of assays for testing compounds that modulate activity of a prostate cancer protein includes the determination of a parameter that is indirectly or directly under the influence of the prostate cancer protein or nucleic acid, e.g., a functional, physical, or chemical effect, such as the ability to decrease prostate cancer. It includes ligand binding activity; cell growth on soft agar; anchorage dependence; contact inhibition and density limitation of growth; cellular proliferation; cellular transformation; growth factor or serum dependence; tumor specific marker levels; invasiveness into Matrigel; tumor growth and metastasis *in vivo*; mRNA and protein expression in cells undergoing metastasis, and other characteristics of prostate cancer cells. "Functional effects" include *in vitro*, *in vivo*, and *ex vivo* activities.

By "determining the functional effect" is meant assaying for a compound that increases or decreases a parameter that is indirectly or directly under the influence of a prostate cancer protein sequence, e.g., functional, enzymatic, physical and chemical effects. Such functional effects can be measured by any means known to those skilled in the art, e.g., changes in spectroscopic characteristics (e.g., fluorescence, absorbance, refractive index), hydrodynamic (e.g., shape), chromatographic, or solubility properties for the protein, measuring inducible markers or transcriptional activation of the prostate cancer protein; measuring binding activity or binding assays, e.g. binding to antibodies or other ligands, and measuring cellular proliferation. Determination of the functional effect of a compound on prostate cancer can also be performed using prostate cancer assays known to those of skill in the art such as an *in vitro* assays, e.g., cell growth on soft agar; anchorage dependence; contact inhibition and density limitation of growth; cellular proliferation; cellular transformation; growth factor or serum dependence; tumor specific marker levels; invasiveness into Matrigel; tumor growth and metastasis *in vivo*; mRNA and protein

expression in cells undergoing metastasis, and other characteristics of prostate cancer cells. The functional effects can be evaluated by many means known to those skilled in the art, e.g., microscopy for quantitative or qualitative measures of alterations in morphological features, measurement of changes in RNA or protein levels for prostate cancer-associated sequences, measurement of RNA stability, identification of downstream or reporter gene expression (CAT, luciferase, β-gal, GFP and the like), e.g., via chemiluminescence, fluorescence, colorimetric reactions, antibody binding, inducible markers, and ligand binding assays.

"Inhibitors", "activators", and "modulators" of prostate cancer polynucleotide and polypeptide sequences are used to refer to activating, inhibitory, or modulating molecules or compounds identified using *in vitro* and *in vivo* assays of prostate cancer polynucleotide and polypeptide sequences. Inhibitors are compounds that, e.g., bind to, partially or totally block activity, decrease, prevent, delay activation, inactivate, desensitize, or down regulate the activity or expression of prostate cancer proteins, e.g., antagonists. Antisense nucleic acids may seem to inhibit expression and subsequent function of the protein. "Activators" are compounds that increase, open, activate, facilitate, enhance activation, sensitize, agonize, or up regulate prostate cancer protein activity. Inhibitors, activators, or modulators also include genetically modified versions of prostate cancer proteins, e.g., versions with altered activity, as well as naturally occurring and synthetic ligands, antagonists, agonists, antibodies, small chemical molecules and the like. Such assays for inhibitors and activators include, e.g., expressing the prostate cancer protein *in vitro*, in cells, or cell membranes, applying putative modulator compounds, and then determining the functional effects on activity, as described above. Activators and inhibitors of prostate cancer can also be identified by incubating prostate cancer cells with the test compound and determining increases or decreases in the expression of 1 or more prostate cancer proteins, e.g., 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 40, 50 or more prostate cancer proteins, such as prostate cancer proteins encoded by the sequences set out in Tables 1-16.

Samples or assays comprising prostate cancer proteins that are treated with a potential activator, inhibitor, or modulator are compared to control samples without the inhibitor, activator, or modulator to examine the extent of inhibition. Control samples (untreated with inhibitors) are assigned a relative protein activity value of 100%. Inhibition of a polypeptide is achieved when the activity value relative to the control is about 80%,

preferably 50%, more preferably 25-0%. Activation of a prostate cancer polypeptide is achieved when the activity value relative to the control (untreated with activators) is 110%, more preferably 150%, more preferably 200-500% (i.e., two to five fold higher relative to the control), more preferably 1000-3000% higher.

The phrase "changes in cell growth" refers to any change in cell growth and proliferation characteristics *in vitro* or *in vivo*, such as formation of foci, anchorage independence, semi-solid or soft agar growth, changes in contact inhibition and density limitation of growth, loss of growth factor or serum requirements, changes in cell morphology, gaining or losing immortalization, gaining or losing tumor specific markers, ability to form or suppress tumors when injected into suitable animal hosts, and/or immortalization of the cell. *See, e.g.,* Freshney, *Culture of Animal Cells a Manual of Basic Technique* pp. 231-241 (3rd ed. 1994).

"Tumor cell" refers to precancerous, cancerous, and normal cells in a tumor.

"Cancer cells," "transformed" cells or "transformation" in tissue culture, refers to spontaneous or induced phenotypic changes that do not necessarily involve the uptake of new genetic material. Although transformation can arise from infection with a transforming virus and incorporation of new genomic DNA, or uptake of exogenous DNA, it can also arise spontaneously or following exposure to a carcinogen, thereby mutating an endogenous gene. Transformation is associated with phenotypic changes, such as immortalization of cells, aberrant growth control, nonmorphological changes, and/or malignancy (*see*, Freshney, *Culture of Animal Cells a Manual of Basic Technique* (3rd ed. 1994)).

"Antibody" refers to a polypeptide comprising a framework region from an immunoglobulin gene or fragments thereof that specifically binds and recognizes an antigen. The recognized immunoglobulin genes include the kappa, lambda, alpha, gamma, delta, epsilon, and mu constant region genes, as well as the myriad immunoglobulin variable region genes. Light chains are classified as either kappa or lambda. Heavy chains are classified as gamma, mu, alpha, delta, or epsilon, which in turn define the immunoglobulin classes, IgG, IgM, IgA, IgD and IgE, respectively. Typically, the antigen-binding region of an antibody or its functional equivalent will be most critical in specificity and affinity of binding. See Paul, *Fundamental Immunology*.

An exemplary immunoglobulin (antibody) structural unit comprises a tetramer. Each tetramer is composed of two identical pairs of polypeptide chains, each pair having one "light" (about 25 kD) and one "heavy" chain (about 50-70 kD). The N-terminus of each chain defines a variable region of about 100 to 110 or more amino acids primarily responsible for antigen recognition. The terms variable light chain (V_L) and variable heavy chain (V_H) refer to these light and heavy chains respectively.

Antibodies exist, e.g., as intact immunoglobulins or as a number of well-characterized fragments produced by digestion with various peptidases. Thus, e.g., pepsin digests an antibody below the disulfide linkages in the hinge region to produce F(ab)'$_2$, a dimer of Fab which itself is a light chain joined to V_H-C_H1 by a disulfide bond. The F(ab)'$_2$ may be reduced under mild conditions to break the disulfide linkage in the hinge region, thereby converting the F(ab)'$_2$ dimer into an Fab' monomer. The Fab' monomer is essentially Fab with part of the hinge region (*see Fundamental Immunology* (Paul ed., 3d ed. 1993). While various antibody fragments are defined in terms of the digestion of an intact antibody, one of skill will appreciate that such fragments may be synthesized *de novo* either chemically or by using recombinant DNA methodology. Thus, the term antibody, as used herein, also includes antibody fragments either produced by the modification of whole antibodies, or those synthesized *de novo* using recombinant DNA methodologies (e.g., single chain Fv) or those identified using phage display libraries (*see, e.g.,* McCafferty *et al.*, *Nature* 348:552-554 (1990))

For preparation of antibodies, e.g., recombinant, monoclonal, or polyclonal antibodies, many technique known in the art can be used (*see, e.g.,* Kohler & Milstein, *Nature* 256:495-497 (1975); Kozbor *et al.*, *Immunology Today* 4:72 (1983); Cole *et al.*, pp. 77-96 in *Monoclonal Antibodies and Cancer Therapy* (1985); Coligan, *Current Protocols in Immunology* (1991); Harlow & Lane, *Antibodies, A Laboratory Manual* (1988); and Goding, *Monoclonal Antibodies: Principles and Practice* (2d ed. 1986)). Techniques for the production of single chain antibodies (U.S. Patent 4,946,778) can be adapted to produce antibodies to polypeptides of this invention. Also, transgenic mice, or other organisms such as other mammals, may be used to express humanized antibodies. Alternatively, phage display technology can be used to identify antibodies and heteromeric Fab fragments that

specifically bind to selected antigens (*see, e.g.*, McCafferty *et al.*, *Nature* 348:552-554 (1990); Marks *et al.*, *Biotechnology* 10:779-783 (1992)).

A "chimeric antibody" is an antibody molecule in which (a) the constant region, or a portion thereof, is altered, replaced or exchanged so that the antigen binding site (variable region) is linked to a constant region of a different or altered class, effector function and/or species, or an entirely different molecule which confers new properties to the chimeric antibody, *e.g.*, an enzyme, toxin, hormone, growth factor, drug, etc.; or (b) the variable region, or a portion thereof, is altered, replaced or exchanged with a variable region having a different or altered antigen specificity.

Identification of prostate cancer-associated sequences

In one aspect, the expression levels of genes are determined in different patient samples for which diagnosis information is desired, to provide expression profiles. An expression profile of a particular sample is essentially a "fingerprint" of the state of the sample; while two states may have any particular gene similarly expressed, the evaluation of a number of genes simultaneously allows the generation of a gene expression profile that is characteristic of the state of the cell. That is, normal tissue (e.g., normal prostate or other tissue) may be distinguished from cancerous or metastatic cancerous tissue of the prostate, or prostate cancer tissue or metastatic prostate cancerous tissue can be compared with tissue samples of prostate and other tissues from surviving cancer patients. By comparing expression profiles of tissue in known different prostate cancer states, information regarding which genes are important (including both up- and down-regulation of genes) in each of these states is obtained.

The identification of sequences that are differentially expressed in prostate cancer versus non-prostate cancer tissue allows the use of this information in a number of ways. For example, a particular treatment regime may be evaluated: does a chemotherapeutic drug act to down-regulate prostate cancer, and thus tumor growth or recurrence, in a particular patient. Similarly, diagnosis and treatment outcomes may be done or confirmed by comparing patient samples with the known expression profiles. Metastatic tissue can also be analyzed to determine the stage of prostate cancer in the tissue. Furthermore, these gene expression profiles (or individual genes) allow screening of drug candidates with an eye to

mimicking or altering a particular expression profile; e.g., screening can be done for drugs that suppress the prostate cancer expression profile. This may be done by making biochips comprising sets of the important prostate cancer genes, which can then be used in these screens. These methods can also be done on the protein basis; that is, protein expression levels of the prostate cancer proteins can be evaluated for diagnostic purposes or to screen candidate agents. In addition, the prostate cancer nucleic acid sequences can be administered for gene therapy purposes, including the administration of antisense nucleic acids, or the prostate cancer proteins (including antibodies and other modulators thereof) administered as therapeutic drugs.

Thus the present invention provides nucleic acid and protein sequences that are differentially expressed in prostate cancer, herein termed "prostate cancer sequences." As outlined below, prostate cancer sequences include those that are up-regulated (i.e., expressed at a higher level) in prostate cancer, as well as those that are down-regulated (i.e., expressed at a lower level). In a preferred embodiment, the prostate cancer sequences are from humans; however, as will be appreciated by those in the art, prostate cancer sequences from other organisms may be useful in animal models of disease and drug evaluation; thus, other prostate cancer sequences are provided, from vertebrates, including mammals, including rodents (rats, mice, hamsters, guinea pigs, etc.), primates, farm animals (including sheep, goats, pigs, cows, horses, etc.) and pets, e.g., (dogs, cats, etc.). Prostate cancer sequences from other organisms may be obtained using the techniques outlined below.

Prostate cancer sequences can include both nucleic acid and amino acid sequences. As will be appreciated by those in the art and is more fully outlined below, prostate cancer nucleic acid sequences are useful in a variety of applications, including diagnostic applications, which will detect naturally occurring nucleic acids, as well as screening applications; e.g., biochips comprising nucleic acid probes or PCR microtiter plates with selected probes to the prostate cancer sequences can be generated.

A prostate cancer sequence can be initially identified by substantial nucleic acid and/or amino acid sequence homology to the prostate cancer sequences outlined herein. Such homology can be based upon the overall nucleic acid or amino acid sequence, and is generally determined as outlined below, using either homology programs or hybridization conditions.

For identifying prostate cancer-associated sequences, the prostate cancer screen typically includes comparing genes identified in different tissues, e.g., normal and cancerous tissues, or tumor tissue samples from patients who have metastatic disease vs. non metastatic tissue. Other suitable tissue comparisons include comparing prostate cancer samples with metastatic cancer samples from other cancers, such as lung, breast, gastrointestinal cancers, ovarian, etc. Samples of different stages of prostate cancer, e.g., survivor tissue, drug resistant states, and tissue undergoing metastasis, are applied to biochips comprising nucleic acid probes. The samples are first microdissected, if applicable, and treated as is known in the art for the preparation of mRNA. Suitable biochips are commercially available, e.g. from Affymetrix. Gene expression profiles as described herein are generated and the data analyzed.

In one embodiment, the genes showing changes in expression as between normal and disease states are compared to genes expressed in other normal tissues, preferably normal prostate, but also including, and not limited to lung, heart, brain, liver, breast, kidney, muscle, colon, small intestine, large intestine, spleen, bone and placenta. In a preferred embodiment, those genes identified during the prostate cancer screen that are expressed in any significant amount in other tissues are removed from the profile, although in some embodiments, this is not necessary. That is, when screening for drugs, it is usually preferable that the target be disease specific, to minimize possible side effects.

In a preferred embodiment, prostate cancer sequences are those that are up-regulated in prostate cancer; that is, the expression of these genes is higher in the prostate cancer tissue as compared to non-cancerous tissue. "Up-regulation" as used herein often means at least about a two-fold change, preferably at least about a three fold change, with at least about five-fold or higher being preferred. All unigene cluster identification numbers and accession numbers herein are for the GenBank sequence database and the sequences of the accession numbers are hereby expressly incorporated by reference. GenBank is known in the art, *see, e.g.,* Benson, DA, *et al.*, Nucleic Acids Research 26:1-7 (1998) and http://www.ncbi.nlm.nih.gov/. Sequences are also available in other databases, e.g., European Molecular Biology Laboratory (EMBL) and DNA Database of Japan (DDBJ).

In another preferred embodiment, prostate cancer sequences are those that are down-regulated in prostate cancer; that is, the expression of these genes is lower in prostate

cancer tissue as compared to non-cancerous tissue (*see, e.g.*, Tables 8, 12 and 14). "Down-regulation" as used herein often means at least about a 1.5-fold change more preferrably a two-fold change, preferably at least about a three fold change, with at least about five-fold or higher being most preferred.

Informatics

The ability to identify genes that are over or under expressed in prostate cancer can additionally provide high-resolution, high-sensitivity datasets which can be used in the areas of diagnostics, therapeutics, drug development, pharmacogenetics, protein structure, biosensor development, and other related areas. For example, the expression profiles can be used in diagnostic or prognostic evaluation of patients with prostate cancer. Or as another example, subcellular toxicological information can be generated to better direct drug structure and activity correlation (*see* Anderson, *Pharmaceutical Proteomics: Targets, Mechanism, and Function*, paper presented at the IBC Proteomics conference, Coronado, CA (June 11-12, 1998)). Subcellular toxicological information can also be utilized in a biological sensor device to predict the likely toxicological effect of chemical exposures and likely tolerable exposure thresholds (*see* U.S. Patent No. 5,811,231). Similar advantages accrue from datasets relevant to other biomolecules and bioactive agents (e.g., nucleic acids, saccharides, lipids, drugs, and the like).

Thus, in another embodiment, the present invention provides a database that includes at least one set of assay data. The data contained in the database is acquired, e.g., using array analysis either singly or in a library format. The database can be in substantially any form in which data can be maintained and transmitted, but is preferably an electronic database. The electronic database of the invention can be maintained on any electronic device allowing for the storage of and access to the database, such as a personal computer, but is preferably distributed on a wide area network, such as the World Wide Web.

The focus of the present section on databases that include peptide sequence data is for clarity of illustration only. It will be apparent to those of skill in the art that similar databases can be assembled for any assay data acquired using an assay of the invention.

The compositions and methods for identifying and/or quantitating the relative and/or absolute abundance of a variety of molecular and macromolecular species from a biological sample undergoing prostate cancer, i.e., the identification of prostate cancer-associated sequences described herein, provide an abundance of information, which can be correlated with pathological conditions, predisposition to disease, drug testing, therapeutic monitoring, gene-disease causal linkages, identification of correlates of immunity and physiological status, among others. Although the data generated from the assays of the invention is suited for manual review and analysis, in a preferred embodiment, prior data processing using high-speed computers is utilized.

An array of methods for indexing and retrieving biomolecular information is known in the art. For example, U.S. Patents 6,023,659 and 5,966,712 disclose a relational database system for storing biomolecular sequence information in a manner that allows sequences to be catalogued and searched according to one or more protein function hierarchies. U.S. Patent 5,953,727 discloses a relational database having sequence records containing information in a format that allows a collection of partial-length DNA sequences to be catalogued and searched according to association with one or more sequencing projects for obtaining full-length sequences from the collection of partial length sequences. U.S. Patent 5,706,498 discloses a gene database retrieval system for making a retrieval of a gene sequence similar to a sequence data item in a gene database based on the degree of similarity between a key sequence and a target sequence. U.S. Patent 5,538,897 discloses a method using mass spectroscopy fragmentation patterns of peptides to identify amino acid sequences in computer databases by comparison of predicted mass spectra with experimentally-derived mass spectra using a closeness-of-fit measure. U.S. Patent 5,926,818 discloses a multi-dimensional database comprising a functionality for multi-dimensional data analysis described as on-line analytical processing (OLAP), which entails the consolidation of projected and actual data according to more than one consolidation path or dimension. U.S. Patent 5,295,261 reports a hybrid database structure in which the fields of each database record are divided into two classes, navigational and informational data, with navigational fields stored in a hierarchical topological map which can be viewed as a tree structure or as the merger of two or more such tree structures.

See also Mount *et al., Bioinformatics* (2001); *Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids* (Durbin *et al.*, eds., 1999); *Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins* (Baxevanis & Oeullette eds., 1998)); Rashidi & Buehler, *Bioinformatics: Basic Applications in Biological Science and Medicine* (1999); *Introduction to Computational Molecular Biology* (Setubal *et al.*, eds 1997); *Bioinformatics: Methods and Protocols* (Misener & Krawetz, eds, 2000); *Bioinformatics: Sequence, Structure, and Databanks: A Practical Approach* (Higgins & Taylor, eds., 2000); Brown, *Bioinformatics: A Biologist's Guide to Biocomputing and the Internet* (2001); Han & Kamber, *Data Mining: Concepts and Techniques* (2000); and Waterman, *Introduction to Computational Biology: Maps, Sequences, and Genomes* (1995).

The present invention provides a computer database comprising a computer and software for storing in computer-retrievable form assay data records cross-tabulated, e.g., with data specifying the source of the target-containing sample from which each sequence specificity record was obtained.

In an exemplary embodiment, at least one of the sources of target-containing sample is from a control tissue sample known to be free of pathological disorders. In a variation, at least one of the sources is a known pathological tissue specimen, e.g., a neoplastic lesion or another tissue specimen to be analyzed for prostate cancer. In another variation, the assay records cross-tabulate one or more of the following parameters for each target species in a sample: (1) a unique identification code, which can include, e.g., a target molecular structure and/or characteristic separation coordinate (e.g., electrophoretic coordinates); (2) sample source; and (3) absolute and/or relative quantity of the target species present in the sample.

The invention also provides for the storage and retrieval of a collection of target data in a computer data storage apparatus, which can include magnetic disks, optical disks, magneto-optical disks, DRAM, SRAM, SGRAM, SDRAM, RDRAM, DDR RAM, magnetic bubble memory devices, and other data storage devices, including CPU registers and on-CPU data storage arrays. Typically, the target data records are stored as a bit pattern in an array of magnetic domains on a magnetizable medium or as an array of charge states or transistor gate states, such as an array of cells in a DRAM device (e.g., each cell comprised of a transistor and a charge storage area, which may be on the transistor). In one embodiment,

the invention provides such storage devices, and computer systems built therewith, comprising a bit pattern encoding a protein expression fingerprint record comprising unique identifiers for at least 10 target data records cross-tabulated with target source.

When the target is a peptide or nucleic acid, the invention preferably provides a method for identifying related peptide or nucleic acid sequences, comprising performing a computerized comparison between a peptide or nucleic acid sequence assay record stored in or retrieved from a computer storage device or database and at least one other sequence. The comparison can include a sequence analysis or comparison algorithm or computer program embodiment thereof (e.g., FASTA, TFASTA, GAP, BESTFIT) and/or the comparison may be of the relative amount of a peptide or nucleic acid sequence in a pool of sequences determined from a polypeptide or nucleic acid sample of a specimen.

The invention also preferably provides a magnetic disk, such as an IBM-compatible (DOS, Windows, Windows95/98/2000, Windows NT, OS/2) or other format (*e.g.*, Linux, SunOS, Solaris, AIX, SCO Unix, VMS, MV, Macintosh, etc.) floppy diskette or hard (fixed, Winchester) disk drive, comprising a bit pattern encoding data from an assay of the invention in a file format suitable for retrieval and processing in a computerized sequence analysis, comparison, or relative quantitation method.

The invention also provides a network, comprising a plurality of computing devices linked via a data link, such as an Ethernet cable (coax or 10BaseT), telephone line, ISDN line, wireless network, optical fiber, or other suitable signal transmission medium, whereby at least one network device (e.g., computer, disk array, etc.) comprises a pattern of magnetic domains (e.g., magnetic disk) and/or charge domains (e.g., an array of DRAM cells) composing a bit pattern encoding data acquired from an assay of the invention.

The invention also provides a method for transmitting assay data that includes generating an electronic signal on an electronic communications device, such as a modem, ISDN terminal adapter, DSL, cable modem, ATM switch, or the like, wherein the signal includes (in native or encrypted format) a bit pattern encoding data from an assay or a database comprising a plurality of assay results obtained by the method of the invention.

In a preferred embodiment, the invention provides a computer system for comparing a query target to a database containing an array of data structures, such as an assay result obtained by the method of the invention, and ranking database targets based on the

degree of identity and gap weight to the target data. A central processor is preferably initialized to load and execute the computer program for alignment and/or comparison of the assay results. Data for a query target is entered into the central processor via an I/O device. Execution of the computer program results in the central processor retrieving the assay data from the data file, which comprises a binary description of an assay result.

The target data or record and the computer program can be transferred to secondary memory, which is typically random access memory (e.g., DRAM, SRAM, SGRAM, or SDRAM). Targets are ranked according to the degree of correspondence between a selected assay characteristic (e.g., binding to a selected affinity moiety) and the same characteristic of the query target and results are output via an I/O device. For example, a central processor can be a conventional computer (e.g., Intel Pentium, PowerPC, Alpha, PA-8000, SPARC, MIPS 4400, MIPS 10000, VAX, etc.); a program can be a commercial or public domain molecular biology software package (e.g., UWGCG Sequence Analysis Software, Darwin); a data file can be an optical or magnetic disk, a data server, a memory device (e.g., DRAM, SRAM, SGRAM, SDRAM, EPROM, bubble memory, flash memory, *etc.*); an I/O device can be a terminal comprising a video display and a keyboard, a modem, an ISDN terminal adapter, an Ethernet port, a punched card reader, a magnetic strip reader, or other suitable I/O device.

The invention also preferably provides the use of a computer system, such as that described above, which comprises: (1) a computer; (2) a stored bit pattern encoding a collection of peptide sequence specificity records obtained by the methods of the invention, which may be stored in the computer; (3) a comparison target, such as a query target; and (4) a program for alignment and comparison, typically with rank-ordering of comparison results on the basis of computed similarity values.

Characteristics of prostate cancer-associated proteins

Prostate cancer proteins of the present invention may be classified as secreted proteins, transmembrane proteins or intracellular proteins. In one embodiment, the prostate cancer protein is an intracellular protein. Intracellular proteins may be found in the cytoplasm and/or in the nucleus. Intracellular proteins are involved in all aspects of cellular function and replication (including, e.g., signaling pathways); aberrant expression of such

proteins often results in unregulated or disregulated cellular processes (*see, e.g., Molecular Biology of the Cell* (Alberts, ed., 3rd ed., 1994). For example, many intracellular proteins have enzymatic activity such as protein kinase activity, protein phosphatase activity, protease activity, nucleotide cyclase activity, polymerase activity and the like. Intracellular proteins also serve as docking proteins that are involved in organizing complexes of proteins, or targeting proteins to various subcellular localizations, and are involved in maintaining the structural integrity of organelles.

An increasingly appreciated concept in characterizing proteins is the presence in the proteins of one or more motifs for which defined functions have been attributed. In addition to the highly conserved sequences found in the enzymatic domain of proteins, highly conserved sequences have been identified in proteins that are involved in protein-protein interaction. For example, Src-homology-2 (SH2) domains bind tyrosine-phosphorylated targets in a sequence dependent manner. PTB domains, which are distinct from SH2 domains, also bind tyrosine phosphorylated targets. SH3 domains bind to proline-rich targets. In addition, PH domains, tetratricopeptide repeats and WD domains to name only a few, have been shown to mediate protein-protein interactions. Some of these may also be involved in binding to phospholipids or other second messengers. As will be appreciated by one of ordinary skill in the art, these motifs can be identified on the basis of primary sequence; thus, an analysis of the sequence of proteins may provide insight into both the enzymatic potential of the molecule and/or molecules with which the protein may associate. One useful database is Pfam (protein families), which is a large collection of multiple sequence alignments and hidden Markov models covering many common protein domains. Versions are available via the internet from Washington University in St. Louis, the Sanger Center in England, and the Karolinska Institute in Sweden (*see, e.g.*, Bateman *et al., Nuc. Acids Res.* 28:263-266 (2000); Sonnhammer *et al., Proteins* 28:405-420 (1997); Bateman *et al., Nuc. Acids Res.* 27:260-262 (1999); and Sonnhammer *et al., Nuc. Acids Res.* 26:320-322-(1998)).

In another embodiment, the prostate cancer sequences are transmembrane proteins. Transmembrane proteins are molecules that span a phospholipid bilayer of a cell. They may have an intracellular domain, an extracellular domain, or both. The intracellular domains of such proteins may have a number of functions including those already described

for intracellular proteins. For example, the intracellular domain may have enzymatic activity and/or may serve as a binding site for additional proteins. Frequently the intracellular domain of transmembrane proteins serves both roles. For example certain receptor tyrosine kinases have both protein kinase activity and SH2 domains. In addition, autophosphorylation of tyrosines on the receptor molecule itself, creates binding sites for additional SH2 domain containing proteins.

Transmembrane proteins may contain from one to many transmembrane domains. For example, receptor tyrosine kinases, certain cytokine receptors, receptor guanylyl cyclases and receptor serine/threonine protein kinases contain a single transmembrane domain. However, various other proteins including channels and adenylyl cyclases contain numerous transmembrane domains. Many important cell surface receptors such as G protein coupled receptors (GPCRs) are classified as "seven transmembrane domain" proteins, as they contain 7 membrane spanning regions. Characteristics of transmembrane domains include approximately 20 consecutive hydrophobic amino acids that may be followed by charged amino acids. Therefore, upon analysis of the amino acid sequence of a particular protein, the localization and number of transmembrane domains within the protein may be predicted (*see, e.g.* PSORT web site http://psort.nibb.ac.jp/). Important transmembrane protein receptors include, but are not limited to the insulin receptor, insulin-like growth factor receptor, human growth hormone receptor, glucose transporters, transferrin receptor, epidermal growth factor receptor, low density lipoprotein receptor, epidermal growth factor receptor, leptin receptor, interleukin receptors, e.g. IL-1 receptor, IL-2 receptor,

The extracellular domains of transmembrane proteins are diverse; however, conserved motifs are found repeatedly among various extracellular domains. Conserved structure and/or functions have been ascribed to different extracellular motifs. Many extracellular domains are involved in binding to other molecules. In one aspect, extracellular domains are found on receptors. Factors that bind the receptor domain include circulating ligands, which may be peptides, proteins, or small molecules such as adenosine and the like. For example, growth factors such as EGF, FGF and PDGF are circulating growth factors that bind to their cognate receptors to initiate a variety of cellular responses. Other factors include cytokines, mitogenic factors, neurotrophic factors and the like. Extracellular domains also

bind to cell-associated molecules. In this respect, they mediate cell-cell interactions. Cell-associated ligands can be tethered to the cell, e.g., via a glycosylphosphatidylinositol (GPI) anchor, or may themselves be transmembrane proteins. Extracellular domains also associate with the extracellular matrix and contribute to the maintenance of the cell structure.

Prostate cancer proteins that are transmembrane are particularly preferred in the present invention as they are readily accessible targets for immunotherapeutics, as are described herein. In addition, as outlined below, transmembrane proteins can be also useful in imaging modalities. Antibodies may be used to label such readily accessible proteins *in situ*. Alternatively, antibodies can also label intracellular proteins, in which case samples are typically permeablized to provide access to intracellular proteins.

It will also be appreciated by those in the art that a transmembrane protein can be made soluble by removing transmembrane sequences, e.g., through recombinant methods. Furthermore, transmembrane proteins that have been made soluble can be made to be secreted through recombinant means by adding an appropriate signal sequence.

In another embodiment, the prostate cancer proteins are secreted proteins; the secretion of which can be either constitutive or regulated. These proteins have a signal peptide or signal sequence that targets the molecule to the secretory pathway. Secreted proteins are involved in numerous physiological events; by virtue of their circulating nature, they serve to transmit signals to various other cell types. The secreted protein may function in an autocrine manner (acting on the cell that secreted the factor), a paracrine manner (acting on cells in close proximity to the cell that secreted the factor) or an endocrine manner (acting on cells at a distance). Thus secreted molecules find use in modulating or altering numerous aspects of physiology. Prostate cancer proteins that are secreted proteins are particularly preferred in the present invention as they serve as good targets for diagnostic markers, *e.g.*, for blood, plasma, serum, or stool tests.

Use of prostate cancer nucleic acids

As described above, prostate cancer sequence is initially identified by substantial nucleic acid and/or amino acid sequence homology or linkage to the prostate cancer sequences outlined herein. Such homology can be based upon the overall nucleic acid or amino acid sequence, and is generally determined as outlined below, using either

homology programs or hybridization conditions. Typically, linked sequences on a mRNA are found on the same molecule.

The prostate cancer nucleic acid sequences of the invention, e.g., the sequences in Tables 1-16, can be fragments of larger genes, i.e., they are nucleic acid segments. "Genes" in this context includes coding regions, non-coding regions, and mixtures of coding and non-coding regions. Accordingly, as will be appreciated by those in the art, using the sequences provided herein, extended sequences, in either direction, of the prostate cancer genes can be obtained, using techniques well known in the art for cloning either longer sequences or the full length sequences; see Ausubel, *et al.*, *supra*. Much can be done by informatics and many sequences can be clustered to include multiple sequences corresponding to a single gene, e.g., systems such as UniGene (see, http://www.ncbi.nlm.nih.gov/UniGene/).

Once the prostate cancer nucleic acid is identified, it can be cloned and, if necessary, its constituent parts recombined to form the entire prostate cancer nucleic acid coding regions or the entire mRNA sequence. Once isolated from its natural source, e.g., contained within a plasmid or other vector or excised therefrom as a linear nucleic acid segment, the recombinant prostate cancer nucleic acid can be further-used as a probe to identify and isolate other prostate cancer nucleic acids, e.g., extended coding regions. It can also be used as a "precursor" nucleic acid to make modified or variant prostate cancer nucleic acids and proteins.

The prostate cancer nucleic acids of the present invention are used in several ways. In a first embodiment, nucleic acid probes to the prostate cancer nucleic acids are made and attached to biochips to be used in screening and diagnostic methods, as outlined below, or for administration, e.g., for gene therapy, vaccine, and/or antisense applications. Alternatively, the prostate cancer nucleic acids that include coding regions of prostate cancer proteins can be put into expression vectors for the expression of prostate cancer proteins, again for screening purposes or for administration to a patient.

In a preferred embodiment, nucleic acid probes to prostate cancer nucleic acids (both the nucleic acid sequences outlined in the figures and/or the complements thereof) are made. The nucleic acid probes attached to the biochip are designed to be substantially complementary to the prostate cancer nucleic acids, *i.e.* the target sequence (either the target

sequence of the sample or to other probe sequences, e.g., in sandwich assays), such that hybridization of the target sequence and the probes of the present invention occurs. As outlined below, this complementarity need not be perfect; there may be any number of base pair mismatches which will interfere with hybridization between the target sequence and the single stranded nucleic acids of the present invention. However, if the number of mutations is so great that no hybridization can occur under even the least stringent of hybridization conditions, the sequence is not a complementary target sequence. Thus, by "substantially complementary" herein is meant that the probes are sufficiently complementary to the target sequences to hybridize under normal reaction conditions, particularly high stringency conditions, as outlined herein.

A nucleic acid probe is generally single stranded but can be partially single and partially double stranded. The strandedness of the probe is dictated by the structure, composition, and properties of the target sequence. In general, the nucleic acid probes range from about 8 to about 100 bases long, with from about 10 to about 80 bases being preferred, and from about 30 to about 50 bases being particularly preferred. That is, generally whole genes are not used. In some embodiments, much longer nucleic acids can be used, up to hundreds of bases.

In a preferred embodiment, more than one probe per sequence is used, with either overlapping probes or probes to different sections of the target being used. That is, two, three, four or more probes, with three being preferred, are used to build in a redundancy for a particular target. The probes can be overlapping (i.e., have some sequence in common), or separate. In some cases, PCR primers may be used to amplify signal for higher sensitivity.

As will be appreciated by those in the art, nucleic acids can be attached or immobilized to a solid support in a wide variety of ways. By "immobilized" and grammatical equivalents herein is meant the association or binding between the nucleic acid probe and the solid support is sufficient to be stable under the conditions of binding, washing, analysis, and removal as outlined below. The binding can typically be covalent or non-covalent. By "non-covalent binding" and grammatical equivalents herein is meant one or more of electrostatic, hydrophilic, and hydrophobic interactions. Included in non-covalent binding is the covalent attachment of a molecule, such as, streptavidin to the support and the non-covalent binding of the biotinylated probe to the streptavidin. By "covalent binding" and grammatical

equivalents herein is meant that the two moieties, the solid support and the probe, are attached by at least one bond, including sigma bonds, pi bonds and coordination bonds. Covalent bonds can be formed directly between the probe and the solid support or can be formed by a cross linker or by inclusion of a specific reactive group on either the solid support or the probe or both molecules. Immobilization may also involve a combination of covalent and non-covalent interactions.

In general, the probes are attached to the biochip in a wide variety of ways, as will be appreciated by those in the art. As described herein, the nucleic acids can either be synthesized first, with subsequent attachment to the biochip, or can be directly synthesized on the biochip.

The biochip comprises a suitable solid substrate. By "substrate" or "solid support" or other grammatical equivalents herein is meant a material that can be modified to contain discrete individual sites appropriate for the attachment or association of the nucleic acid probes and is amenable to at least one detection method. As will be appreciated by those in the art, the number of possible substrates are very large, and include, but are not limited to, glass and modified or functionalized glass, plastics (including acrylics, polystyrene and copolymers of styrene and other materials, polypropylene, polyethylene, polybutylene, polyurethanes, TeflonJ, etc.), polysaccharides, nylon or nitrocellulose, resins, silica or silica-based materials including silicon and modified silicon, carbon, metals, inorganic glasses, plastics, etc. In general, the substrates allow optical detection and do not appreciably fluoresce. A preferred substrate is described in copending application entitled Reusable Low Fluorescent Plastic Biochip, U.S. Application Serial No. 09/270,214, filed March 15, 1999, herein incorporated by reference in its entirety.

Generally the substrate is planar, although as will be appreciated by those in the art, other configurations of substrates may be used as well. For example, the probes may be placed on the inside surface of a tube, for flow-through sample analysis to minimize sample volume. Similarly, the substrate may be flexible, such as a flexible foam, including closed cell foams made of particular plastics.

In a preferred embodiment, the surface of the biochip and the probe may be derivatized with chemical functional groups for subsequent attachment of the two. Thus, e.g., the biochip is derivatized with a chemical functional group including, but not limited to,

amino groups, carboxy groups, oxo groups and thiol groups, with amino groups being particularly preferred. Using these functional groups, the probes can be attached using functional groups on the probes. For example, nucleic acids containing amino groups can be attached to surfaces comprising amino groups, e.g. using linkers as are known in the art; e.g., homo-or hetero-bifunctional linkers as are well known (*see* 1994 Pierce Chemical Company catalog, technical section on cross-linkers, pages 155-200). In addition, in some cases, additional linkers, such as alkyl groups (including substituted and heteroalkyl groups) may be used.

In this embodiment, oligonucleotides are synthesized as is known in the art, and then attached to the surface of the solid support. As will be appreciated by those skilled in the art, either the 5' or 3' terminus may be attached to the solid support, or attachment may be via an internal nucleoside.

In another embodiment, the immobilization to the solid support may be very strong, yet non-covalent. For example, biotinylated oligonucleotides can be made, which bind to surfaces covalently coated with streptavidin, resulting in attachment.

Alternatively, the oligonucleotides may be synthesized on the surface, as is known in the art. For example, photoactivation techniques utilizing photopolymerization compounds and techniques are used. In a preferred embodiment, the nucleic acids can be synthesized in situ, using well known photolithographic techniques, such as those described in WO 95/25116; WO 95/35505; U.S. Patent Nos. 5,700,637 and 5,445,934; and references cited within, all of which are expressly incorporated by reference; these methods of attachment form the basis of the Affimetrix GeneChip™ technology.

Often, amplification-based assays are performed to measure the expression level of prostate cancer-associated sequences. These assays are typically performed in conjunction with reverse transcription. In such assays, a prostate cancer-associated nucleic acid sequence acts as a template in an amplification reaction (e.g., Polymerase Chain Reaction, or PCR). In a quantitative amplification, the amount of amplification product will be proportional to the amount of template in the original sample. Comparison to appropriate controls provides a measure of the amount of prostate cancer-associated RNA. Methods of quantitative amplification are well known to those of skill in the art. Detailed protocols for

quantitative PCR are provided, e.g., in Innis *et al.*, *PCR Protocols, A Guide to Methods and Applications* (1990).

In some embodiments, a TaqMan based assay is used to measure expression. TaqMan based assays use a fluorogenic oligonucleotide probe that contains a 5' fluorescent dye and a 3' quenching agent. The probe hybridizes to a PCR product, but cannot itself be extended due to a blocking agent at the 3' end. When the PCR product is amplified in subsequent cycles, the 5' nuclease activity of the polymerase, *e.g.*, AmpliTaq, results in the cleavage of the TaqMan probe. This cleavage separates the 5' fluorescent dye and the 3' quenching agent, thereby resulting in an increase in fluorescence as a function of amplification (*see*, e.g., literature provided by Perkin-Elmer, *e.g.*, www2.perkin-elmer.com).

Other suitable amplification methods include, but are not limited to, ligase chain reaction (LCR) (*see* Wu & Wallace, *Genomics* 4:560 (1989), Landegren *et al.*, *Science* 241:1077 (1988), and Barringer *et al.*, *Gene* 89:117 (1990)), transcription amplification (Kwoh *et al.*, *Proc. Natl. Acad. Sci. USA* 86:1173 (1989)), self-sustained sequence replication (Guatelli *et al.*, *Proc. Nat. Acad. Sci. USA* 87:1874 (1990)), dot PCR, and linker adapter PCR, etc.

Expression of prostate cancer proteins from nucleic acids

In a preferred embodiment, prostate cancer nucleic acids, e.g., encoding prostate cancer proteins are used to make a variety of expression vectors to express prostate cancer proteins which can then be used in screening assays, as described below. Expression vectors and recombinant DNA technology are well known to those of skill in the art (*see, e.g.*, Ausubel, *supra*, and *Gene Expression Systems* (Fernandez & Hoeffler, eds, 1999)) and are used to express proteins. The expression vectors may be either self-replicating extrachromosomal vectors or vectors which integrate into a host genome. Generally, these expression vectors include transcriptional and translational regulatory nucleic acid operably linked to the nucleic acid encoding the prostate cancer protein. The term "control sequences" refers to DNA sequences used for the expression of an operably linked coding sequence in a particular host organism. Control sequences that are suitable for prokaryotes, e.g., include a promoter, optionally an operator sequence, and a ribosome binding site. Eukaryotic cells are known to utilize promoters, polyadenylation signals, and enhancers.

Nucleic acid is "operably linked" when it is placed into a functional relationship with another nucleic acid sequence. For example, DNA for a presequence or secretory leader is operably linked to DNA for a polypeptide if it is expressed as a preprotein that participates in the secretion of the polypeptide; a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; or a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation. Generally, "operably linked" means that the DNA sequences being linked are contiguous, and, in the case of a secretory leader, contiguous and in reading phase. However, enhancers do not have to be contiguous. Linking is typically accomplished by ligation at convenient restriction sites. If such sites do not exist, synthetic oligonucleotide adaptors or linkers are used in accordance with conventional practice. Transcriptional and translational regulatory nucleic acid will generally be appropriate to the host cell used to express the prostate cancer protein. Numerous types of appropriate expression vectors, and suitable regulatory sequences are known in the art for a variety of host cells.

In general, transcriptional and translational regulatory sequences may include, but are not limited to, promoter sequences, ribosomal binding sites, transcriptional start and stop sequences, translational start and stop sequences, and enhancer or activator sequences. In a preferred embodiment, the regulatory sequences include a promoter and transcriptional start and stop sequences.

Promoter sequences encode either constitutive or inducible promoters. The promoters may be either naturally occurring promoters or hybrid promoters. Hybrid promoters, which combine elements of more than one promoter, are also known in the art, and are useful in the present invention.

In addition, an expression vector may comprise additional elements. For example, the expression vector may have two replication systems, thus allowing it to be maintained in two organisms, e.g. in mammalian or insect cells for expression and in a procaryotic host for cloning and amplification. Furthermore, for integrating expression vectors, the expression vector contains at least one sequence homologous to the host cell genome, and preferably two homologous sequences which flank the expression construct. The integrating vector may be directed to a specific locus in the host cell by selecting the

appropriate homologous sequence for inclusion in the vector. Constructs for integrating vectors are well known in the art (e.g., Fernandez & Hoeffler, *supra*).

In addition, in a preferred embodiment, the expression vector contains a selectable marker gene to allow the selection of transformed host cells. Selection genes are well known in the art and will vary with the host cell used.

The prostate cancer proteins of the present invention are produced by culturing a host cell transformed with an expression vector containing nucleic acid encoding a prostate cancer protein, under the appropriate conditions to induce or cause expression of the prostate cancer protein. Conditions appropriate for prostate cancer protein expression will vary with the choice of the expression vector and the host cell, and will be easily ascertained by one skilled in the art through routine experimentation or optimization. For example, the use of constitutive promoters in the expression vector will require optimizing the growth and proliferation of the host cell, while the use of an inducible promoter requires the appropriate growth conditions for induction. In addition, in some embodiments, the timing of the harvest is important. For example, the baculoviral systems used in insect cell expression are lytic viruses, and thus harvest time selection can be crucial for product yield.

Appropriate host cells include yeast, bacteria, archaebacteria, fungi, and insect and animal cells, including mammalian cells. Of particular interest are *Saccharomyces cerevisiae* and other yeasts, *E. coli, Bacillus subtilis*, Sf9 cells, C129 cells, 293 cells, *Neurospora*, BHK, CHO, COS, HeLa cells, HUVEC (human umbilical vein endothelial cells), THP1 cells (a macrophage cell line) and various other human cells and cell lines.

In a preferred embodiment, the prostate cancer proteins are expressed in mammalian cells. Mammalian expression systems are also known in the art, and include retroviral and adenoviral systems. One expression vector system is a retroviral vector system such as is generally described in PCT/US97/01019 and PCT/US97/01048, both of which are hereby expressly incorporated by reference. Of particular use as mammalian promoters are the promoters from mammalian viral genes, since the viral genes are often highly expressed and have a broad host range. Examples include the SV40 early promoter, mouse mammary tumor virus LTR promoter, adenovirus major late promoter, herpes simplex virus promoter, and the CMV promoter (*see, e.g.,* Fernandez & Hoeffler, *supra*). Typically, transcription termination and polyadenylation sequences recognized by mammalian cells are regulatory

regions located 3' to the translation stop codon and thus, together with the promoter elements, flank the coding sequence. Examples of transcription terminator and polyadenlyation signals include those derived form SV40.

The methods of introducing exogenous nucleic acid into mammalian hosts, as well as other hosts, is well known in the art, and will vary with the host cell used. Techniques include dextran-mediated transfection, calcium phosphate precipitation, polybrene mediated transfection, protoplast fusion, electroporation, viral infection, encapsulation of the polynucleotide(s) in liposomes, and direct microinjection of the DNA into nuclei.

In a preferred embodiment, prostate cancer proteins are expressed in bacterial systems. Bacterial expression systems are well known in the art. Promoters from bacteriophage may also be used and are known in the art. In addition, synthetic promoters and hybrid promoters are also useful; e.g., the tac promoter is a hybrid of the trp and lac promoter sequences. Furthermore, a bacterial promoter can include naturally occurring promoters of non-bacterial origin that have the ability to bind bacterial RNA polymerase and initiate transcription. In addition to a functioning promoter sequence, an efficient ribosome binding site is desirable. The expression vector may also include a signal peptide sequence that provides for secretion of the prostate cancer protein in bacteria. The protein is either secreted into the growth media (gram-positive bacteria) or into the periplasmic space, located between the inner and outer membrane of the cell (gram-negative bacteria). The bacterial expression vector may also include a selectable marker gene to allow for the selection of bacterial strains that have been transformed. Suitable selection genes include genes which render the bacteria resistant to drugs such as ampicillin, chloramphenicol, erythromycin, kanamycin, neomycin and tetracycline. Selectable markers also include biosynthetic genes, such as those in the histidine, tryptophan and leucine biosynthetic pathways. These components are assembled into expression vectors. Expression vectors for bacteria are well known in the art, and include vectors for *Bacillus subtilis, E. coli, Streptococcus cremoris*, and *Streptococcus lividans*, among others (e.g., Fernandez & Hoeffler, *supra*). The bacterial expression vectors are transformed into bacterial host cells using techniques well known in the art, such as calcium chloride treatment, electroporation, and others.

In one embodiment, prostate cancer proteins are produced in insect cells. Expression vectors for the transformation of insect cells, and in particular, baculovirus-based expression vectors, are well known in the art.

In a preferred embodiment, prostate cancer protein is produced in yeast cells. Yeast expression systems are well known in the art, and include expression vectors for *Saccharomyces cerevisiae, Candida albicans* and *C. maltosa, Hansenula polymorpha, Kluyveromyces fragilis* and *K. lactis, Pichia guillerimondii and P. pastoris, Schizosaccharomyces pombe*, and *Yarrowia lipolytica.*

The prostate cancer protein may also be made as a fusion protein, using techniques well known in the art. Thus, e.g., for the creation of monoclonal antibodies, if the desired epitope is small, the prostate cancer protein may be fused to a carrier protein to form an immunogen. Alternatively, the prostate cancer protein may be made as a fusion protein to increase expression, or for other reasons. For example, when the prostate cancer protein is a prostate cancer peptide, the nucleic acid encoding the peptide may be linked to other nucleic acid for expression purposes.

In a preferred embodiment, the prostate cancer protein is purified or isolated after expression. Prostate cancer proteins may be isolated or purified in a variety of ways known to those skilled in the art depending on what other components are present in the sample. Standard purification methods include electrophoretic, molecular, immunological and chromatographic techniques, including ion exchange, hydrophobic, affinity, and reverse-phase HPLC chromatography, and chromatofocusing. For example, the prostate cancer protein may be purified using a standard anti-prostate cancer protein antibody column. Ultrafiltration and diafiltration techniques, in conjunction with protein concentration, are also useful. For general guidance in suitable purification techniques, see Scopes, *Protein Purification* (1982). The degree of purification necessary will vary depending on the use of the prostate cancer protein. In some instances no purification will be necessary.

Once expressed and purified if necessary, the prostate cancer proteins and nucleic acids are useful in a number of applications. They may be used as immunoselection reagents, as vaccine reagents, as screening agents, etc.

Variants of prostate cancer proteins

In one embodiment, the prostate cancer proteins are derivative or variant prostate cancer proteins as compared to the wild-type sequence. That is, as outlined more fully below, the derivative prostate cancer peptide will often contain at least one amino acid substitution, deletion or insertion, with amino acid substitutions being particularly preferred. The amino acid substitution, insertion or deletion may occur at any residue within the prostate cancer peptide.

Also included within one embodiment of prostate cancer proteins of the present invention are amino acid sequence variants. These variants typically fall into one or more of three classes: substitutional, insertional or deletional variants. These variants ordinarily are prepared by site specific mutagenesis of nucleotides in the DNA encoding the prostate cancer protein, using cassette or PCR mutagenesis or other techniques well known in the art, to produce DNA encoding the variant, and thereafter expressing the DNA in recombinant cell culture as outlined above. However, variant prostate cancer protein fragments having up to about 100-150 residues may be prepared by in vitro synthesis using established techniques. Amino acid sequence variants are characterized by the predetermined nature of the variation, a feature that sets them apart from naturally occurring allelic or interspecies variation of the prostate cancer protein amino acid sequence. The variants typically exhibit the same qualitative biological activity as the naturally occurring analogue, although variants can also be selected which have modified characteristics as will be more fully outlined below.

While the site or region for introducing an amino acid sequence variation is predetermined, the mutation per se need not be predetermined. For example, in order to optimize the performance of a mutation at a given site, random mutagenesis may be conducted at the target codon or region and the expressed prostate cancer variants screened for the optimal combination of desired activity. Techniques for making substitution mutations at predetermined sites in DNA having a known sequence are well known, e.g., M13 primer mutagenesis and PCR mutagenesis. Screening of the mutants is done using assays of prostate cancer protein activities.

Amino acid substitutions are typically of single residues; insertions usually will be on the order of from about 1 to 20 amino acids, although considerably larger

insertions may be tolerated. Deletions range from about 1 to about 20 residues, although in some cases deletions may be much larger.

Substitutions, deletions, insertions or any combination thereof may be used to arrive at a final derivative. Generally these changes are done on a few amino acids to minimize the alteration of the molecule. However, larger changes may be tolerated in certain circumstances. When small alterations in the characteristics of the prostate cancer protein are desired, substitutions are generally made in accordance with the amino acid substitution relationships provided in the definition section.

The variants typically exhibit the same qualitative biological activity and will elicit the same immune response as the naturally-occurring analog, although variants also are selected to modify the characteristics of the prostate cancer proteins as needed. Alternatively, the variant may be designed such that the biological activity of the prostate cancer protein is altered. For example, glycosylation sites may be altered or removed.

Substantial changes in function or immunological identity are made by selecting substitutions that are less conservative than those described above. For example, substitutions may be made which more significantly affect: the structure of the polypeptide backbone in the area of the alteration, for example the alpha-helical or beta-sheet structure; the charge or hydrophobicity of the molecule at the target site; or the bulk of the side chain. The substitutions which in general are expected to produce the greatest changes in the polypeptide's properties are those in which (a) a hydrophilic residue, e.g. seryl or threonyl is substituted for (or by) a hydrophobic residue, e.g. leucyl, isoleucyl, phenylalanyl, valyl or alanyl; (b) a cysteine or proline is substituted for (or by) any other residue; (c) a residue having an electropositive side chain, e.g. lysyl, arginyl, or histidyl, is substituted for (or by) an electronegative residue, e.g. glutamyl or aspartyl; or (d) a residue having a bulky side chain, e.g. phenylalanine, is substituted for (or by) one not having a side chain, e.g. glycine.

Covalent modifications of prostate cancer polypeptides are included within the scope of this invention. One type of covalent modification includes reacting targeted amino acid residues of a prostate cancer polypeptide with an organic derivatizing agent that is capable of reacting with selected side chains or the N-or C-terminal residues of a prostate cancer polypeptide. Derivatization with bifunctional agents is useful, for instance, for crosslinking prostate cancer polypeptides to a water-insoluble support matrix or surface for

use in the method for purifying anti-prostate cancer polypeptide antibodies or screening assays, as is more fully described below. Commonly used crosslinking agents include, e.g., 1,1-bis(diazoacetyl)-2-phenylethane, glutaraldehyde, N-hydroxysuccinimide esters, e.g., esters with 4-azidosalicylic acid, homobifunctional imidoesters, including disuccinimidyl esters such as 3,3'-dithiobis(succinimidylpropionate), bifunctional maleimides such as bis-N-maleimido-1,8-octane and agents such as methyl-3-((p-azidophenyl)dithio)propioimidate.

Other modifications include deamidation of glutaminyl and asparaginyl residues to the corresponding glutamyl and aspartyl residues, respectively, hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of seryl, threonyl or tyrosyl residues, methylation of the amino groups of the lysine, arginine, and histidine side chains (Creighton, *Proteins: Structure and Molecular Properties*, pp. 79-86 (1983)), acetylation of the N-terminal amine, and amidation of any C-terminal carboxyl group.

Another type of covalent modification of the prostate cancer polypeptide included within the scope of this invention comprises altering the native glycosylation pattern of the polypeptide. "Altering the native glycosylation pattern" is intended for purposes herein to mean deleting one or more carbohydrate moieties found in native sequence prostate cancer polypeptide, and/or adding one or more glycosylation sites that are not present in the native sequence prostate cancer polypeptide. Glycosylation patterns can be altered in many ways. For example the use of different cell types to express prostate cancer-associated sequences can result in different glycosylation patterns.

Addition of glycosylation sites to prostate cancer polypeptides may also be accomplished by altering the amino acid sequence thereof. The alteration may be made, e.g., by the addition of, or substitution by, one or more serine or threonine residues to the native sequence prostate cancer polypeptide (for O-linked glycosylation sites). The prostate cancer amino acid sequence may optionally be altered through changes at the DNA level, particularly by mutating the DNA encoding the prostate cancer polypeptide at preselected bases such that codons are generated that will translate into the desired amino acids.

Another means of increasing the number of carbohydrate moieties on the prostate cancer polypeptide is by chemical or enzymatic coupling of glycosides to the polypeptide. Such methods are described in the art, e.g., in WO 87/05330, and in Aplin & Wriston, *CRC Crit. Rev. Biochem.*, pp. 259-306 (1981).

Removal of carbohydrate moieties present on the prostate cancer polypeptide may be accomplished chemically or enzymatically or by mutational substitution of codons encoding for amino acid residues that serve as targets for glycosylation. Chemical deglycosylation techniques are known in the art and described, for instance, by Hakimuddin, *et al.*, *Arch. Biochem. Biophys.*, 259:52 (1987) and by Edge *et al.*, *Anal. Biochem.*, 118:131 (1981). Enzymatic cleavage of carbohydrate moieties on polypeptides can be achieved by the use of a variety of endo-and exo-glycosidases as described by Thotakura *et al.*, *Meth. Enzymol.*, 138:350 (1987).

Another type of covalent modification of prostate cancer comprises linking the prostate cancer polypeptide to one of a variety of nonproteinaceous polymers, e.g., polyethylene glycol, polypropylene glycol, or polyoxyalkylenes, in the manner set forth in U.S. Patent Nos. 4,640,835; 4,496,689; 4,301,144; 4,670,417; 4,791,192 or 4,179,337.

Prostate cancer polypeptides of the present invention may also be modified in a way to form chimeric molecules comprising a prostate cancer polypeptide fused to another, heterologous polypeptide or amino acid sequence. In one embodiment, such a chimeric molecule comprises a fusion of a prostate cancer polypeptide with a tag polypeptide which provides an epitope to which an anti-tag antibody can selectively bind. The epitope tag is generally placed at the amino-or carboxyl-terminus of the prostate cancer polypeptide. The presence of such epitope-tagged forms of a prostate cancer polypeptide can be detected using an antibody against the tag polypeptide. Also, provision of the epitope tag enables the prostate cancer polypeptide to be readily purified by affinity purification using an anti-tag antibody or another type of affinity matrix that binds to the epitope tag. In an alternative embodiment, the chimeric molecule may comprise a fusion of a prostate cancer polypeptide with an immunoglobulin or a particular region of an immunoglobulin. For a bivalent form of the chimeric molecule, such a fusion could be to the Fc region of an IgG molecule.

Various tag polypeptides and their respective antibodies are well known in the art. Examples include poly-histidine (poly-his) or poly-histidine-glycine (poly-his-gly) tags; HIS6 and metal chelation tags, the flu HA tag polypeptide and its antibody 12CA5 (Field *et al.*, *Mol. Cell. Biol.* 8:2159-2165 (1988)); the c-myc tag and the 8F9, 3C7, 6E10, G4, B7 and 9E10 antibodies thereto (Evan *et al.*, *Molecular and Cellular Biology* 5:3610-3616 (1985)); and the Herpes Simplex virus glycoprotein D (gD) tag and its antibody (*Paborsky et al.*,

Protein Engineering 3(6):547-553 (1990)). Other tag polypeptides include the Flag-peptide (*Hopp et al., BioTechnology* 6:1204-1210 (1988)); the KT3 epitope peptide (Martin *et al., Science* 255:192-194 (1992)); tubulin epitope peptide (Skinner *et al., J. Biol. Chem.* 266:15163-15166 (1991)); and the T7 gene 10 protein peptide tag (Lutz-Freyermuth *et al., Proc. Natl. Acad. Sci. USA* 87:6393-6397 (1990)).

Also included are other prostate cancer proteins of the prostate cancer family, and prostate cancer proteins from other organisms, which are cloned and expressed as outlined below. Thus, probe or degenerate polymerase chain reaction (PCR) primer sequences may be used to find other related prostate cancer proteins from humans or other organisms. As will be appreciated by those in the art, particularly useful probe and/or PCR primer sequences include the unique areas of the prostate cancer nucleic acid sequence. As is generally known in the art, preferred PCR primers are from about 15 to about 35 nucleotides in length, with from about 20 to about 30 being preferred, and may contain inosine as needed. The conditions for the PCR reaction are well known in the art (e.g., Innis, PCR Protocols, *supra*).

Antibodies to prostate cancer proteins

In a preferred embodiment, when the prostate cancer protein is to be used to generate antibodies, e.g., for immunotherapy or immunodiagnosis, the prostate cancer protein should share at least one epitope or determinant with the full length protein. By "epitope" or "determinant" herein is typically meant a portion of a protein which will generate and/or bind an antibody or T-cell receptor in the context of MHC. Thus, in most instances, antibodies made to a smaller prostate cancer protein will be able to bind to the full-length protein, particularly linear epitopes. In a preferred embodiment; the epitope is unique; that is, antibodies generated to a unique epitope show little or no cross-reactivity.

Methods of preparing polyclonal antibodies are known to the skilled artisan (e.g., Coligan, *supra*; and Harlow & Lane, *supra*). Polyclonal antibodies can be raised in a mammal, e.g., by one or more injections of an immunizing agent and, if desired, an adjuvant. Typically, the immunizing agent and/or adjuvant will be injected in the mammal by multiple subcutaneous or intraperitoneal injections. The immunizing agent may include a protein encoded by a nucleic acid of the figures or fragment thereof or a fusion protein thereof. It

may be useful to conjugate the immunizing agent to a protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. Examples of adjuvants which may be employed include Freund's complete adjuvant and MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate). The immunization protocol may be selected by one skilled in the art without undue experimentation.

The antibodies may, alternatively, be monoclonal antibodies. Monoclonal antibodies may be prepared using hybridoma methods, such as those described by Kohler & Milstein, *Nature* 256:495 (1975). In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes may be immunized in vitro. The immunizing agent will typically include a polypeptide encoded by a nucleic acid of Tables 1-16 fragment thereof, or a fusion protein thereof. Generally, either peripheral blood lymphocytes ("PBLs") are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell (Goding, *Monoclonal Antibodies: Principles and Practice*, pp. 59-103 (1986)). Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells may be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

In one embodiment, the antibodies are bispecific antibodies. Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens or that have binding specificities for two epitopes on the same antigen. In one embodiment, one of the binding specificities is for a

protein encoded by a nucleic acid Tables 1-16 or a fragment thereof, the other one is for any other antigen, and preferably for a cell-surface protein or receptor or receptor subunit, preferably one that is tumor specific. Alternatively, tetramer-type technology may create multivalent reagents.

In a preferred embodiment, the antibodies to prostate cancer protein are capable of reducing or eliminating a biological function of a prostate cancer protein, as is described below. That is, the addition of anti-prostate cancer protein antibodies (either polyclonal or preferably monoclonal) to prostate cancer tissue (or cells containing prostate cancer) may reduce or eliminate the prostate cancer. Generally, at least a 25% decrease in activity, growth, size or the like is preferred, with at least about 50% being particularly preferred and about a 95-100% decrease being especially preferred.

In a preferred embodiment the antibodies to the prostate cancer proteins are humanized antibodies (e.g., Xenerex Biosciences, Mederex, Inc., Abgenix, Inc., Protein Design Labs,Inc.) Humanized forms of non-human (e.g., murine) antibodies are chimeric molecules of immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')2 or other antigen-binding subsequences of antibodies) which contain minimal sequence derived from non-human immunoglobulin. Humanized antibodies include human immunoglobulins (recipient antibody) in which residues from a complementary determining region (CDR) of the recipient are replaced by residues from a CDR of a non-human species (donor antibody) such as mouse, rat or rabbit having the desired specificity, affinity and capacity. In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies may also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, a humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the framework (FR) regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin (Jones *et al.*, *Nature* 321:522-525 (1986); Riechmann *et al.*, *Nature* 332:323-329 (1988); and Presta, *Curr. Op. Struct. Biol.* 2:593-596 (1992)). Humanization

can be essentially performed following the method of Winter and co-workers (Jones *et al.*, *Nature* 321:522-525 (1986); Riechmann *et al.*, *Nature* 332:323-327 (1988); Verhoeyen *et al.*, *Science* 239:1534-1536 (1988)), by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. Accordingly, such humanized antibodies are chimeric antibodies (U.S. Patent No. 4,816,567), wherein substantially less than an intact human variable domain has been substituted by the corresponding sequence from a non-human species.

Human antibodies can also be produced using various techniques known in the art, including phage display libraries (Hoogenboom & Winter, *J. Mol. Biol.* 227:381 (1991); Marks *et al.*, *J. Mol. Biol.* 222:581 (1991)). The techniques of Cole *et al.* and Boerner *et al.* are also available for the preparation of human monoclonal antibodies (Cole *et al.*, *Monoclonal Antibodies and Cancer Therapy,* p. 77 (1985) and Boerner *et al.*, *J. Immunol.* 147(1):86-95 (1991)). Similarly, human antibodies can be made by introducing of human immunoglobulin loci into transgenic animals, e.g., mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely resembles that seen in humans in all respects, including gene rearrangement, assembly, and antibody repertoire. This approach is described, e.g., in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in the following scientific publications: Marks *et al.*, *Bio/Technology* 10:779-783 (1992); Lonberg *et al.*, *Nature* 368:856-859 (1994); Morrison, *Nature* 368:812-13 (1994); Fishwild *et al.*, *Nature Biotechnology* 14:845-51 (1996); Neuberger, *Nature Biotechnology* 14:826 (1996); Lonberg & Huszar, *Intern. Rev. Immunol.* 13:65-93 (1995).

By immunotherapy is meant treatment of prostate cancer with an antibody raised against prostate cancer proteins. As used herein, immunotherapy can be passive or active. Passive immunotherapy as defined herein is the passive transfer of antibody to a recipient (patient). Active immunization is the induction of antibody and/or T-cell responses in a recipient (patient). Induction of an immune response is the result of providing the recipient with an antigen to which antibodies are raised. As appreciated by one of ordinary skill in the art, the antigen may be provided by injecting a polypeptide against which antibodies are desired to be raised into a recipient, or contacting the recipient with a nucleic

acid capable of expressing the antigen and under conditions for expression of the antigen, leading to an immune response.

In a preferred embodiment the prostate cancer proteins against which antibodies are raised are secreted proteins as described above. Without being bound by theory, antibodies used for treatment, bind and prevent the secreted protein from binding to its receptor, thereby inactivating the secreted prostate cancer protein.

In another preferred embodiment, the prostate cancer protein to which antibodies are raised is a transmembrane protein. Without being bound by theory, antibodies used for treatment, bind the extracellular domain of the prostate cancer protein and prevent it from binding to other proteins, such as circulating ligands or cell-associated molecules. The antibody may cause down-regulation of the transmembrane prostate cancer protein. As will be appreciated by one of ordinary skill in the art, the antibody may be a competitive, non-competitive or uncompetitive inhibitor of protein binding to the extracellular domain of the prostate cancer protein. The antibody is also an antagonist of the prostate cancer protein. Further, the antibody prevents activation of the transmembrane prostate cancer protein. In one aspect, when the antibody prevents the binding of other molecules to the prostate cancer protein, the antibody prevents growth of the cell. The antibody may also be used to target or sensitize the cell to cytotoxic agents, including, but not limited to TNF-α, TNF-β, IL-1, INF-γ and IL-2, or chemotherapeutic agents including 5FU, vinblastine, actinomycin D, cisplatin, methotrexate, and the like. In some instances the antibody belongs to a sub-type that activates serum complement when complexed with the transmembrane protein thereby mediating cytotoxicity or antigen-dependent cytotoxicity (ADCC). Thus, prostate cancer is treated by administering to a patient antibodies directed against the transmembrane prostate cancer protein. Antibody-labeling may activate a co-toxin, localize a toxin payload, or otherwise provide means to locally ablate cells.

In another preferred embodiment, the antibody is conjugated to an effector moiety. The effector moiety can be any number of molecules, including labelling moieties such as radioactive labels or fluorescent labels, or can be a therapeutic moiety. In one aspect the therapeutic moiety is a small molecule that modulates the activity of the prostate cancer protein. In another aspect the therapeutic moiety modulates the activity of molecules associated with or in close proximity to the prostate cancer protein. The therapeutic moiety

may inhibit enzymatic activity such as protease or collagenase or protein kinase activity associated with prostate cancer.

In a preferred embodiment, the therapeutic moiety can also be a cytotoxic agent. In this method, targeting the cytotoxic agent to prostate cancer tissue or cells, results in a reduction in the number of afflicted cells, thereby reducing symptoms associated with prostate cancer. Cytotoxic agents are numerous and varied and include, but are not limited to, cytotoxic drugs or toxins or active fragments of such toxins. Suitable toxins and their corresponding fragments include diphtheria A chain, exotoxin A chain, ricin A chain, abrin A chain, curcin, crotin, phenomycin, enomycin and the like. Cytotoxic agents also include radiochemicals made by conjugating radioisotopes to antibodies raised against prostate cancer proteins, or binding of a radionuclide to a chelating agent that has been covalently attached to the antibody. Targeting the therapeutic moiety to transmembrane prostate cancer proteins not only serves to increase the local concentration of therapeutic moiety in the prostate cancer afflicted area, but also serves to reduce deleterious side effects that may be associated with the therapeutic moiety.

In another preferred embodiment, the prostate cancer protein against which the antibodies are raised is an intracellular protein. In this case, the antibody may be conjugated to a protein which facilitates entry into the cell. In one case, the antibody enters the cell by endocytosis. In another embodiment, a nucleic acid encoding the antibody is administered to the individual or cell. Moreover, wherein the prostate cancer protein can be targeted within a cell, i.e., the nucleus, an antibody thereto contains a signal for that target localization, i.e., a nuclear localization signal.

The prostate cancer antibodies of the invention specifically bind to prostate cancer proteins. By "specifically bind" herein is meant that the antibodies bind to the protein with a K_d of at least about 0.1 mM, more usually at least about 1 μM, preferably at least about 0.1 μM or better, and most preferably, 0.01 μM or better. Selectivity of binding is also important.

Detection of prostate cancer sequence for diagnostic and therapeutic applications

In one aspect, the RNA expression levels of genes are determined for different cellular states in the prostate cancer phenotype. Expression levels of genes in normal tissue

(i.e., not undergoing prostate cancer) and in prostate cancer tissue (and in some cases, for varying severities of prostate cancer that relate to prognosis, as outlined below) are evaluated to provide expression profiles. An expression profile of a particular cell state or point of development is essentially a "fingerprint" of the state. While two states may have any particular gene similarly expressed, the evaluation of a number of genes simultaneously allows the generation of a gene expression profile that is reflective of the state of the cell. By comparing expression profiles of cells in different states, information regarding which genes are important (including both up- and down-regulation of genes) in each of these states is obtained. Then, diagnosis may be performed or confirmed to determine whether a tissue sample has the gene expression profile of normal or cancerous tissue. This will provide for molecular diagnosis of related conditions.

"Differential expression," or grammatical equivalents as used herein, refers to qualitative or quantitative differences in the temporal and/or cellular gene expression patterns within and among cells and tissue. Thus, a differentially expressed gene can qualitatively have its expression altered, including an activation or inactivation, in, e.g., normal versus prostate cancer tissue. Genes may be turned on or turned off in a particular state, relative to another state thus permitting comparison of two or more states. A qualitatively regulated gene will exhibit an expression pattern within a state or cell type which is detectable by standard techniques. Some genes will be expressed in one state or cell type, but not in both. Alternatively, the difference in expression may be quantitative, e.g., in that expression is increased or decreased; i.e., gene expression is either upregulated, resulting in an increased amount of transcript, or downregulated, resulting in a decreased amount of transcript. The degree to which expression differs need only be large enough to quantify via standard characterization techniques as outlined below, such as by use of Affymetrix GeneChip™ expression arrays, Lockhart, *Nature Biotechnology* 14:1675-1680 (1996), hereby expressly incorporated by reference. Other techniques include, but are not limited to, quantitative reverse transcriptase PCR, northern analysis and RNase protection. As outlined above, preferably the change in expression (i.e., upregulation or downregulation) is at least about 50%, more preferably at least about 100%, more preferably at least about 150%, more preferably at least about 200%, with from 300 to at least 1000% being especially preferred.

Evaluation may be at the gene transcript, or the protein level. The amount of gene expression may be monitored using nucleic acid probes to the DNA or RNA equivalent of the gene transcript, and the quantification of gene expression levels, or, alternatively, the final gene product itself (protein) can be monitored, e.g., with antibodies to the prostate cancer protein and standard immunoassays (ELISAs, etc.) or other techniques, including mass spectroscopy assays, 2D gel electrophoresis assays, etc. Proteins corresponding to prostate cancer genes, i.e., those identified as being important in a prostate cancer phenotype, can be evaluated in a prostate cancer diagnostic test.

In a preferred embodiment, gene expression monitoring is performed simultaneously on a number of genes. Multiple protein expression monitoring can be performed as well. Similarly, these assays may be performed on an individual basis as well.

In this embodiment, the prostate cancer nucleic acid probes are attached to biochips as outlined herein for the detection and quantification of prostate cancer sequences in a particular cell. The assays are further described below in the example. PCR techniques can be used to provide greater sensitivity.

In a preferred embodiment nucleic acids encoding the prostate cancer protein are detected. Although DNA or RNA encoding the prostate cancer protein may be detected, of particular interest are methods wherein an mRNA encoding a prostate cancer protein is detected. Probes to detect mRNA can be a nucleotide/deoxynucleotide probe that is complementary to and hybridizes with the mRNA and includes, but is not limited to, oligonucleotides, cDNA or RNA. Probes also should contain a detectable label, as defined herein. In one method the mRNA is detected after immobilizing the nucleic acid to be examined on a solid support such as nylon membranes and hybridizing the probe with the sample. Following washing to remove the non-specifically bound probe, the label is detected. In another method detection of the mRNA is performed in situ. In this method permeabilized cells or tissue samples are contacted with a detectably labeled nucleic acid probe for sufficient time to allow the probe to hybridize with the target mRNA. Following washing to remove the non-specifically bound probe, the label is detected. For example a digoxygenin labeled riboprobe (RNA probe) that is complementary to the mRNA encoding a prostate cancer protein is detected by binding the digoxygenin with an anti-digoxygenin

secondary antibody and developed with nitro blue tetrazolium and 5-bromo-4-chloro-3-indoyl phosphate.

In a preferred embodiment, various proteins from the three classes of proteins as described herein (secreted, transmembrane or intracellular proteins) are used in diagnostic assays. The prostate cancer proteins, antibodies, nucleic acids, modified proteins and cells containing prostate cancer sequences are used in diagnostic assays. This can be performed on an individual gene or corresponding polypeptide level. In a preferred embodiment, the expression profiles are used, preferably in conjunction with high throughput screening techniques to allow monitoring for expression profile genes and/or corresponding polypeptides.

As described and defined herein, prostate cancer proteins, including intracellular, transmembrane or secreted proteins, find use as markers of prostate cancer. Detection of these proteins in putative prostate cancer tissue allows for detection or diagnosis of prostate cancer. In one embodiment, antibodies are used to detect prostate cancer proteins. A preferred method separates proteins from a sample by electrophoresis on a gel (typically a denaturing and reducing protein gel, but may be another type of gel, including isoelectric focusing gels and the like). Following separation of proteins, the prostate cancer protein is detected, e.g., by immunoblotting with antibodies raised against the prostate cancer protein. Methods of immunoblotting are well known to those of ordinary skill in the art.

In another preferred method, antibodies to the prostate cancer protein find use in *in situ* imaging techniques, e.g., in histology (e.g., *Methods in Cell Biology: Antibodies in Cell Biology*, volume 37 (Asai, ed. 1993)). In this method cells are contacted with from one to many antibodies to the prostate cancer protein(s). Following washing to remove non-specific antibody binding, the presence of the antibody or antibodies is detected. In one embodiment the antibody is detected by incubating with a secondary antibody that contains a detectable label. In another method the primary antibody to the prostate cancer protein(s) contains a detectable label, e.g. an enzyme marker that can act on a substrate. In another preferred embodiment each one of multiple primary antibodies contains a distinct and detectable label. This method finds particular use in simultaneous screening for a plurality of prostate cancer proteins. As will be appreciated by one of ordinary skill in the art, many other histological imaging techniques are also provided by the invention.

In a preferred embodiment the label is detected in a fluorometer which has the ability to detect and distinguish emissions of different wavelengths. In addition, a fluorescence activated cell sorter (FACS) can be used in the method.

In another preferred embodiment, antibodies find use in diagnosing prostate cancer from blood, serum, plasma, stool, and other samples. Such samples, therefore, are useful as samples to be probed or tested for the presence of prostate cancer proteins. Antibodies can be used to detect a prostate cancer protein by previously described immunoassay techniques including ELISA, immunoblotting (western blotting), immunoprecipitation, BIACORE technology and the like. Conversely, the presence of antibodies may indicate an immune response against an endogenous prostate cancer protein.

In a preferred embodiment, *in situ* hybridization of labeled prostate cancer nucleic acid probes to tissue arrays is done. For example, arrays of tissue samples, including prostate cancer tissue and/or normal tissue, are made. *In situ* hybridization (*see, e.g.,* Ausubel, *supra*) is then performed. When comparing the fingerprints between an individual and a standard, the skilled artisan can make a diagnosis, a prognosis, or a prediction based on the findings. It is further understood that the genes which indicate the diagnosis may differ from those which indicate the prognosis and molecular profiling of the condition of the cells may lead to distinctions between responsive or refractory conditions or may be predictive of outcomes.

In a preferred embodiment, the prostate cancer proteins, antibodies, nucleic acids, modified proteins and cells containing prostate cancer sequences are used in prognosis assays. As above, gene expression profiles can be generated that correlate to prostate cancer, in terms of long term prognosis. Again, this may be done on either a protein or gene level, with the use of genes being preferred. As above, prostate cancer probes may be attached to biochips for the detection and quantification of prostate cancer sequences in a tissue or patient. The assays proceed as outlined above for diagnosis. PCR method may provide more sensitive and accurate quantification.

Assays for therapeutic compounds

In a preferred embodiment members of the proteins, nucleic acids, and antibodies as described herein are used in drug screening assays. The prostate cancer

proteins, antibodies, nucleic acids, modified proteins and cells containing prostate cancer sequences are used in drug screening assays or by evaluating the effect of drug candidates on a "gene expression profile" or expression profile of polypeptides. In a preferred embodiment, the expression profiles are used, preferably in conjunction with high throughput screening techniques to allow monitoring for expression profile genes after treatment with a candidate agent (e.g., Zlokarnik, *et al.*, *Science* 279:84-8 (1998); Heid, *Genome Res* 6:986-94, 1996).

In a preferred embodiment, the prostate cancer proteins, antibodies, nucleic acids, modified proteins and cells containing the native or modified prostate cancer proteins are used in screening assays. That is, the present invention provides novel methods for screening for compositions which modulate the prostate cancer phenotype or an identified physiological function of a prostate cancer protein. As above, this can be done on an individual gene level or by evaluating the effect of drug candidates on a "gene expression profile". In a preferred embodiment, the expression profiles are used, preferably in conjunction with high throughput screening techniques to allow monitoring for expression profile genes after treatment with a candidate agent, see Zlokarnik, *supra.*

Having identified the differentially expressed genes herein, a variety of assays may be executed. In a preferred embodiment, assays may be run on an individual gene or protein level. That is, having identified a particular gene as up regulated in prostate cancer, test compounds can be screened for the ability to modulate gene expression or for binding to the prostate cancer protein. "Modulation" thus includes both an increase and a decrease in gene expression. The preferred amount of modulation will depend on the original change of the gene expression in normal versus tissue undergoing prostate cancer, with changes of at least 10%, preferably 50%, more preferably 100-300%, and in some embodiments 300-1000% or greater. Thus, if a gene exhibits a 4-fold increase in prostate cancer tissue compared to normal tissue, a decrease of about four-fold is often desired; similarly, a 10-fold decrease in prostate cancer tissue compared to normal tissue often provides a target value of a 10-fold increase in expression to be induced by the test compound.

The amount of gene expression may be monitored using nucleic acid probes and the quantification of gene expression levels, or, alternatively, the gene product itself can be monitored, e.g., through the use of antibodies to the prostate cancer protein and standard

immunoassays. Proteomics and separation techniques may also allow quantification of expression.

In a preferred embodiment, gene expression or protein monitoring of a number of entities, i.e., an expression profile, is monitored simultaneously. Such profiles will typically involve a plurality of those entities described herein..

In this embodiment, the prostate cancer nucleic acid probes are attached to biochips as outlined herein for the detection and quantification of prostate cancer sequences in a particular cell. Alternatively, PCR may be used. Thus, a series, e.g., of microtiter plate, may be used with dispensed primers in desired wells. A PCR reaction can then be performed and analyzed for each well.

Expression monitoring can be performed to identify compounds that modify the expression of one or more prostate cancer-associated sequences, e.g., a polynucleotide sequence set out in Tables 1-16. Generally, in a preferred embodiment, a test modulator is added to the cells prior to analysis. Moreover, screens are also provided to identify agents that modulate prostate cancer, modulate prostate cancer proteins, bind to a prostate cancer protein, or interfere with the binding of a prostate cancer protein and an antibody or other binding partner.

The term "test compound" or "drug candidate" or "modulator" or grammatical equivalents as used herein describes any molecule, e.g., protein, oligopeptide, small organic molecule, polysaccharide, polynucleotide, etc., to be tested for the capacity to directly or indirectly alter the prostate cancer phenotype or the expression of a prostate cancer sequence, e.g., a nucleic acid or protein sequence. In preferred embodiments, modulators alter expression profiles, or expression profile nucleic acids or proteins provided herein. In one embodiment, the modulator suppresses a prostate cancer phenotype, e.g. to a normal tissue fingerprint. In another embodiment, a modulator induced a prostate cancer phenotype. Generally, a plurality of assay mixtures are run in parallel with different agent concentrations to obtain a differential response to the various concentrations. Typically, one of these concentrations serves as a negative control, i.e., at zero concentration or below the level of detection.

Drug candidates encompass numerous chemical classes, though typically they are organic molecules, preferably small organic compounds having a molecular weight of

more than 100 and less than about 2,500 daltons. Preferred small molecules are less than 2000, or less than 1500 or less than 1000 or less than 500 D. Candidate agents comprise functional groups necessary for structural interaction with proteins, particularly hydrogen bonding, and typically include at least an amine, carbonyl, hydroxyl or carboxyl group, preferably at least two of the functional chemical groups. The candidate agents often comprise cyclical carbon or heterocyclic structures and/or aromatic or polyaromatic structures substituted with one or more of the above functional groups. Candidate agents are also found among biomolecules including peptides, saccharides, fatty acids, steroids, purines, pyrimidines, derivatives, structural analogs or combinations thereof. Particularly preferred are peptides.

In one aspect, a modulator will neutralize the effect of a prostate cancer protein. By "neutralize" is meant that activity of a protein is inhibited or blocked and the consequent effect on the cell.

In certain embodiments, combinatorial libraries of potential modulators will be screened for an ability to bind to a prostate cancer polypeptide or to modulate activity. Conventionally, new chemical entities with useful properties are generated by identifying a chemical compound (called a "lead compound") with some desirable property or activity, e.g., inhibiting activity, creating variants of the lead compound, and evaluating the property and activity of those variant compounds. Often, high throughput screening (HTS) methods are employed for such an analysis.

In one preferred embodiment, high throughput screening methods involve providing a library containing a large number of potential therapeutic compounds (candidate compounds). Such "combinatorial chemical libraries" are then screened in one or more assays to identify those library members (particular chemical species or subclasses) that display a desired characteristic activity. The compounds thus identified can serve as conventional "lead compounds" or can themselves be used as potential or actual therapeutics.

A combinatorial chemical library is a collection of diverse chemical compounds generated by either chemical synthesis or biological synthesis by combining a number of chemical "building blocks" such as reagents. For example, a linear combinatorial chemical library, such as a polypeptide (e.g., mutein) library, is formed by combining a set of chemical building blocks called amino acids in every possible way for a given compound

length (i.e., the number of amino acids in a polypeptide compound). Millions of chemical compounds can be synthesized through such combinatorial mixing of chemical building blocks (Gallop *et al., J. Med. Chem.* 37(9):1233-1251 (1994)).

Preparation and screening of combinatorial chemical libraries is well known to those of skill in the art. Such combinatorial chemical libraries include, but are not limited to, peptide libraries (*see, e.g.,* U.S. Patent No. 5,010,175, Furka, *Pept. Prot. Res.* 37:487-493 (1991), Houghton *et al., Nature,* 354:84-88 (1991)), peptoids (PCT Publication No WO 91/19735), encoded peptides (PCT Publication WO 93/20242), random bio-oligomers (PCT Publication WO 92/00091), benzodiazepines (U.S. Pat. No. 5,288,514), diversomers such as hydantoins, benzodiazepines and dipeptides (Hobbs *et al., Proc. Nat. Acad. Sci. USA* 90:6909-6913 (1993)), vinylogous polypeptides (Hagihara *et al., J. Amer. Chem. Soc.* 114:6568 (1992)), nonpeptidal peptidomimetics with a Beta-D-Glucose scaffolding (Hirschmann *et al., J. Amer. Chem. Soc.* 114:9217-9218 (1992)), analogous organic syntheses of small compound libraries (Chen *et al., J. Amer. Chem.* Soc. 116:2661 (1994)), oligocarbamates (Cho, *et al., Science* 261:1303 (1993)), and/or peptidyl phosphonates (Campbell *et al., J. Org. Chem.* 59:658 (1994)). *See, generally,* Gordon *et al., J. Med. Chem.* 37:1385 (1994), nucleic acid libraries (*see, e.g.,* Strategene, Corp.), peptide nucleic acid libraries (*see, e.g.,* U.S. Patent 5,539,083), antibody libraries (*see, e.g.,* Vaughn *et al., Nature Biotechnology* 14(3):309-314 (1996), and PCT/US96/10287), carbohydrate libraries (*see, e.g.,* Liang *et al., Science* 274:1520-1522 (1996), and U.S. Patent No. 5,593,853), and small organic molecule libraries (*see, e.g.,* benzodiazepines, Baum, C&EN, Jan 18, page 33 (1993); isoprenoids, U.S. Patent No. 5,569,588; thiazolidinones and metathiazanones, U.S. Patent No. 5,549,974; pyrrolidines, U.S. Patent Nos. 5,525,735 and 5,519,134; morpholino compounds, U.S. Patent No. 5,506,337; benzodiazepines, U.S. Patent No. 5,288,514; and the like).

Devices for the preparation of combinatorial libraries are commercially available (*see, e.g.,* 357 MPS, 390 MPS, Advanced Chem Tech, Louisville KY, Symphony, Rainin, Woburn, MA, 433A Applied Biosystems, Foster City, CA, 9050 Plus, Millipore, Bedford, MA).

A number of well known robotic systems have also been developed for solution phase chemistries. These systems include automated workstations like the automated synthesis apparatus developed by Takeda Chemical Industries, LTD. (Osaka,

Japan) and many robotic systems utilizing robotic arms (Zymate II, Zymark Corporation, Hopkinton, Mass.; Orca, Hewlett-Packard, Palo Alto, Calif.), which mimic the manual synthetic operations performed by a chemist. Any of the above devices are suitable for use with the present invention. The nature and implementation of modifications to these devices (if any) so that they can operate as discussed herein will be apparent to persons skilled in the relevant art. In addition, numerous combinatorial libraries are themselves commercially available (*see, e.g.,* ComGenex, Princeton, N.J., Asinex, Moscow, Ru, Tripos, Inc., St. Louis, MO, ChemStar, Ltd, Moscow, RU, 3D Pharmaceuticals, Exton, PA, Martek Biosciences, Columbia, MD, *etc.*).

The assays to identify modulators are amenable to high throughput screening. Preferred assays thus detect enhancement or inhibition of prostate cancer gene transcription, inhibition or enhancement of polypeptide expression, and inhibition or enhancement of polypeptide activity.

High throughput assays for the presence, absence, quantification, or other properties of particular nucleic acids or protein products are well known to those of skill in the art. Similarly, binding assays and reporter gene assays are similarly well known. Thus, e.g., U.S. Patent No. 5,559,410 discloses high throughput screening methods for proteins, U.S. Patent No. 5,585,639 discloses high throughput screening methods for nucleic acid binding (i.e., in arrays), while U.S. Patent Nos. 5,576,220 and 5,541,061 disclose high throughput methods of screening for ligand/antibody binding.

In addition, high throughput screening systems are commercially available (*see, e.g.,* Zymark Corp., Hopkinton, MA; Air Technical Industries, Mentor, OH; Beckman Instruments, Inc. Fullerton, CA; Precision Systems, Inc., Natick, MA, *etc.*). These systems typically automate entire procedures, including all sample and reagent pipetting, liquid dispensing, timed incubations, and final readings of the microplate in detector(s) appropriate for the assay. These configurable systems provide high throughput and rapid start up as well as a high degree of flexibility and customization. The manufacturers of such systems provide detailed protocols for various high throughput systems. Thus, e.g., Zymark Corp. provides technical bulletins describing screening systems for detecting the modulation of gene transcription, ligand binding, and the like.

In one embodiment, modulators are proteins, often naturally occurring proteins or fragments of naturally occurring proteins. Thus, *e.g.*, cellular extracts containing proteins, or random or directed digests of proteinaceous cellular extracts, may be used. In this way libraries of proteins may be made for screening in the methods of the invention. Particularly preferred in this embodiment are libraries of bacterial, fungal, viral, and mammalian proteins, with the latter being preferred, and human proteins being especially preferred. Particularly useful test compound will be directed to the class of proteins to which the target belongs, *e.g.*, substrates for enzymes or ligands and receptors.

In a preferred embodiment, modulators are peptides of from about 5 to about 30 amino acids, with from about 5 to about 20 amino acids being preferred, and from about 7 to about 15 being particularly preferred. The peptides may be digests of naturally occurring proteins as is outlined above, random peptides, or "biased" random peptides. By "randomized" or grammatical equivalents herein is meant that each nucleic acid and peptide consists of essentially random nucleotides and amino acids, respectively. Since generally these random peptides (or nucleic acids, discussed below) are chemically synthesized, they may incorporate any nucleotide or amino acid at any position. The synthetic process can be designed to generate randomized proteins or nucleic acids, to allow the formation of all or most of the possible combinations over the length of the sequence, thus forming a library of randomized candidate bioactive proteinaceous agents.

In one embodiment, the library is fully randomized, with no sequence preferences or constants at any position. In a preferred embodiment, the library is biased. That is, some positions within the sequence are either held constant, or are selected from a limited number of possibilities. For example, in a preferred embodiment, the nucleotides or amino acid residues are randomized within a defined class, e.g., of hydrophobic amino acids, hydrophilic residues, sterically biased (either small or large) residues, towards the creation of nucleic acid binding domains, the creation of cysteines, for cross-linking, prolines for SH-3 domains, serines, threonines, tyrosines or histidines for phosphorylation sites, etc., or to purines, etc.

Modulators of prostate cancer can also be nucleic acids, as defined below. As described above generally for proteins, nucleic acid modulating agents may be naturally occurring nucleic acids, random nucleic acids, or "biased" random nucleic acids. For

example, digests of procaryotic or eucaryotic genomes may be used as is outlined above for proteins.

In certain embodiments, the activity of a prostate cancer-associated protein is down-regulated, or entirely inhibited, by the use of antisense polynucleotide, *i.e.*, a nucleic acid complementary to, and which can preferably hybridize specifically to, a coding mRNA nucleic acid sequence, e.g., a prostate cancer protein mRNA, or a subsequence thereof. Binding of the antisense polynucleotide to the mRNA reduces the translation and/or stability of the mRNA.

In the context of this invention, antisense polynucleotides can comprise naturally-occurring nucleotides, or synthetic species formed from naturally-occurring subunits or their close homologs. Antisense polynucleotides may also have altered sugar moieties or inter-sugar linkages. Exemplary among these are the phosphorothioate and other sulfur containing species which are known for use in the art. Analogs are comprehended by this invention so long as they function effectively to hybridize with the prostate cancer protein mRNA. *See, e.g.,* Isis Pharmaceuticals, Carlsbad, CA; Sequitor, Inc., Natick, MA.

Such antisense polynucleotides can readily be synthesized using recombinant means, or can be synthesized *in vitro*. Equipment for such synthesis is sold by several vendors, including Applied Biosystems. The preparation of other oligonucleotides such as phosphorothioates and alkylated derivatives is also well known to those of skill in the art.

Antisense molecules as used herein include antisense or sense oligonucleotides. Sense oligonucleotides can, e.g., be employed to block transcription by binding to the anti-sense strand. The antisense and sense oligonucleotide comprise a single-stranded nucleic acid sequence (either RNA or DNA) capable of binding to target mRNA (sense) or DNA (antisense) sequences for prostate cancer molecules. Antisense or sense oligonucleotides, according to the present invention, comprise a fragment generally at least about 14 nucleotides, preferably from about 14 to 30 nucleotides. The ability to derive an antisense or a sense oligonucleotide, based upon a cDNA sequence encoding a given protein is described in, e.g., Stein & Cohen (*Cancer Res.* 48:2659 (1988 and van der Krol *et al.* (*BioTechniques* 6:958 (1988)).

In addition to antisense polynucleotides, ribozymes can be used to target and inhibit transcription of prostate cancer-associated nucleotide sequences. A ribozyme is an

RNA molecule that catalytically cleaves other RNA molecules. Different kinds of ribozymes have been described, including group I ribozymes, hammerhead ribozymes, hairpin ribozymes, RNase P, and axhead ribozymes (*see, e.g.,* Castanotto *et al.*, *Adv. in Pharmacology* 25: 289-317 (1994) for a general review of the properties of different ribozymes).

The general features of hairpin ribozymes are described, e.g., in Hampel *et al.*, *Nucl. Acids Res.* 18:299-304 (1990); European Patent Publication No. 0 360 257; U.S. Patent No. 5,254,678. Methods of preparing are well known to those of skill in the art (*see, e.g.,* WO 94/26877; Ojwang *et al.*, *Proc. Natl. Acad. Sci.* USA 90:6340-6344 (1993); Yamada *et al.*, *Human Gene Therapy* 1:39-45 (1994); Leavitt *et al.*, *Proc. Natl. Acad. Sci. USA* 92:699-703 (1995); Leavitt *et al.*, *Human Gene Therapy* 5:1151-120 (1994); and Yamada *et al.*, *Virology* 205: 121-126 (1994)).

Polynucleotide modulators of prostate cancer may be introduced into a cell containing the target nucleotide sequence by formation of a conjugate with a ligand binding molecule, as described in WO 91/04753. Suitable ligand binding molecules include, but are not limited to, cell surface receptors, growth factors, other cytokines, or other ligands that bind to cell surface receptors. Preferably, conjugation of the ligand binding molecule does not substantially interfere with the ability of the ligand binding molecule to bind to its corresponding molecule or receptor, or block entry of the sense or antisense oligonucleotide or its conjugated version into the cell. Alternatively, a polynucleotide modulator of prostate cancer may be introduced into a cell containing the target nucleic acid sequence, e.g., by formation of an polynucleotide-lipid complex, as described in WO 90/10448. It is understood that the use of antisense molecules or knock out and knock in models may also be used in screening assays as discussed above, in addition to methods of treatment.

As noted above, gene expression monitoring is conveniently used to test candidate modultors (*e.g.*, protein, nucleic acid or small molecule). After the candidate agent has been added and the cells allowed to incubate for some period of time, the sample containing a target sequence to be analyzed is added to the biochip. If required, the target sequence is prepared using known techniques. For example, the sample may be treated to lyse the cells, using known lysis buffers, electroporation, etc., with purification and/or amplification such as PCR performed as appropriate. For example, an *in vitro* transcription

with labels covalently attached to the nucleotides is performed. Generally, the nucleic acids are labeled with biotin-FITC or PE, or with cy3 or cy5.

In a preferred embodiment, the target sequence is labeled with, e.g., a fluorescent, a chemiluminescent, a chemical, or a radioactive signal, to provide a means of detecting the target sequence's specific binding to a probe. The label also can be an enzyme, such as, alkaline phosphatase or horseradish peroxidase, which when provided with an appropriate substrate produces a product that can be detected. Alternatively, the label can be a labeled compound or small molecule, such as an enzyme inhibitor, that binds but is not catalyzed or altered by the enzyme. The label also can be a moiety or compound, such as, an epitope tag or biotin which specifically binds to streptavidin. For the example of biotin, the streptavidin is labeled as described above, thereby, providing a detectable signal for the bound target sequence. Unbound labeled streptavidin is typically removed prior to analysis.

As will be appreciated by those in the art, these assays can be direct hybridization assays or can comprise "sandwich assays", which include the use of multiple probes, as is generally outlined in U.S. Patent Nos. 5,681,702, 5,597,909, 5,545,730, 5,594,117, 5,591,584, 5,571,670, 5,580,731, 5,571,670, 5,591,584, 5,624,802, 5,635,352, 5,594,118, 5,359,100, 5,124,246 and 5,681,697, all of which are hereby incorporated by reference. In this embodiment, in general, the target nucleic acid is prepared as outlined above, and then added to the biochip comprising a plurality of nucleic acid probes, under conditions that allow the formation of a hybridization complex.

A variety of hybridization conditions may be used in the present invention, including high, moderate and low stringency conditions as outlined above. The assays are generally run under stringency conditions which allows formation of the label probe hybridization complex only in the presence of target. Stringency can be controlled by altering a step parameter that is a thermodynamic variable, including, but not limited to, temperature, formamide concentration, salt concentration, chaotropic salt concentration pH, organic solvent concentration, etc.

These parameters may also be used to control non-specific binding, as is generally outlined in U.S. Patent No. 5,681,697. Thus it may be desirable to perform certain steps at higher stringency conditions to reduce non-specific binding.

The reactions outlined herein may be accomplished in a variety of ways. Components of the reaction may be added simultaneously, or sequentially, in different orders, with preferred embodiments outlined below. In addition, the reaction may include a variety of other reagents. These include salts, buffers, neutral proteins, e.g. albumin, detergents, *etc.* which may be used to facilitate optimal hybridization and detection, and/or reduce non-specific or background interactions. Reagents that otherwise improve the efficiency of the assay, such as protease inhibitors, nuclease inhibitors, anti-microbial agents, *etc.*, may also be used as appropriate, depending on the sample preparation methods and purity of the target.

The assay data are analyzed to determine the expression levels, and changes in expression levels as between states, of individual genes, forming a gene expression profile.

Screens are performed to identify modulators of the prostate cancer phenotype. In one embodiment, screening is performed to identify modulators that can induce or suppress a particular expression profile, thus preferably generating the associated phenotype. In another embodiment, *e.g.*, for diagnostic applications, having identified differentially expressed genes important in a particular state, screens can be performed to identify modulators that alter expression of individual genes. In an another embodiment, screening is performed to identify modulators that alter a biological function of the expression product of a differentially expressed gene. Again, having identified the importance of a gene in a particular state, screens are performed to identify agents that bind and/or modulate the biological activity of the gene product.

In addition screens can be done for genes that are induced in response to a candidate agent. After identifying a modulator based upon its ability to suppress a prostate cancer expression pattern leading to a normal expression pattern, or to modulate a single prostate cancer gene expression profile so as to mimic the expression of the gene from normal tissue, a screen as described above can be performed to identify genes that are specifically modulated in response to the agent. Comparing expression profiles between normal tissue and agent treated prostate cancer tissue reveals genes that are not expressed in normal tissue or prostate cancer tissue, but are expressed in agent treated tissue. These agent-specific sequences can be identified and used by methods described herein for prostate cancer genes or proteins. In particular these sequences and the proteins they encode find use in marking or identifying agent treated cells. In addition, antibodies can be raised against the

agent induced proteins and used to target novel therapeutics to the treated prostate cancer tissue sample.

Thus, in one embodiment, a test compound is administered to a population of prostate cancer cells, that have an associated prostate cancer expression profile. By "administration" or "contacting" herein is meant that the candidate agent is added to the cells in such a manner as to allow the agent to act upon the cell, whether by uptake and intracellular action, or by action at the cell surface. In some embodiments, nucleic acid encoding a proteinaceous candidate agent (i.e., a peptide) may be put into a viral construct such as an adenoviral or retroviral construct, and added to the cell, such that expression of the peptide agent is accomplished, e.g., PCT US97/01019. Regulatable gene therapy systems can also be used.

Once the test compound has been administered to the cells, the cells can be washed if desired and are allowed to incubate under preferably physiological conditions for some period of time. The cells are then harvested and a new gene expression profile is generated, as outlined herein.

Thus, e.g., prostate cancer tissue may be screened for agents that modulate, e.g., induce or suppress the prostate cancer phenotype. A change in at least one gene, preferably many, of the expression profile indicates that the agent has an effect on prostate cancer activity. By defining such a signature for the prostate cancer phenotype, screens for new drugs that alter the phenotype can be devised. With this approach, the drug target need not be known and need not be represented in the original expression screening platform, nor does the level of transcript for the target protein need to change.

In a preferred embodiment, as outlined above, screens may be done on individual genes and gene products (proteins). That is, having identified a particular differentially expressed gene as important in a particular state, screening of modulators of either the expression of the gene or the gene product itself can be done. The gene products of differentially expressed genes are sometimes referred to herein as "prostate cancer proteins" or a "prostate cancer modulatory protein". The prostate cancer modulatory protein may be a fragment, or alternatively, be the full length protein to the fragment encoded by the nucleic acids of Tables 1-16. Preferably, the prostate cancer modulatory protein is a fragment. In a preferred embodiment, the prostate cancer amino acid sequence which is used to determine

sequence identity or similarity is encoded by a nucleic acid of Tables 1-16. In another embodiment, the sequences are naturally occurring allelic variants of a protein encoded by a nucleic acid of Tables 1-16. In another embodiment, the sequences are sequence variants as further described herein.

Preferably, the prostate cancer modulatory protein is a fragment of approximately 14 to 24 amino acids long. More preferably the fragment is a soluble fragment. Preferably, the fragment includes a non-transmembrane region. In a preferred embodiment, the fragment has an N-terminal Cys to aid in solubility. In one embodiment, the C-terminus of the fragment is kept as a free acid and the N-terminus is a free amine to aid in coupling, i.e., to cysteine.

In one embodiment the prostate cancer proteins are conjugated to an immunogenic agent as discussed herein. In one embodiment the prostate cancer protein is conjugated to BSA.

Measurements of prostate cancer polypeptide activity, or of prostate cancer or the prostate cancer phenotype can be performed using a variety of assays. For example, the effects of the test compounds upon the function of the prostate cancer polypeptides can be measured by examining parameters described above. A suitable physiological change that affects activity can be used to assess the influence of a test compound on the polypeptides of this invention. When the functional consequences are determined using intact cells or animals, one can also measure a variety of effects such as, in the case of prostate cancer associated with tumors, tumor growth, tumor metastasis, neovascularization, hormone release, transcriptional changes to both known and uncharacterized genetic markers (e.g., northern blots), changes in cell metabolism such as cell growth or pH changes, and changes in intracellular second messengers such as cGMP. In the assays of the invention, mammalian prostate cancer polypeptide is typically used, e.g., mouse, preferably human.

Assays to identify compounds with modulating activity can be performed *in vitro*. For example, a prostate cancer polypeptide is first contacted with a potential modulator and incubated for a suitable amount of time, e.g., from 0.5 to 48 hours. In one embodiment, the prostate cancer polypeptide levels are determined *in vitro* by measuring the level of protein or mRNA. The level of protein is measured using immunoassays such as western blotting, ELISA and the like with an antibody that selectively binds to the prostate cancer

polypeptide or a fragment thereof. For measurement of mRNA, amplification, e.g., using PCR, LCR, or hybridization assays, e.g., northern hybridization, RNAse protection, dot blotting, are preferred. The level of protein or mRNA is detected using directly or indirectly labeled detection agents, e.g., fluorescently or radioactively labeled nucleic acids, radioactively or enzymatically labeled antibodies, and the like, as described herein.

Alternatively, a reporter gene system can be devised using the prostate cancer protein promoter operably linked to a reporter gene such as luciferase, green fluorescent protein, CAT, or β-gal. The reporter construct is typically transfected into a cell. After treatment with a potential modulator, the amount of reporter gene transcription, translation, or activity is measured according to standard techniques known to those of skill in the art.

In a preferred embodiment, as outlined above, screens may be done on individual genes and gene products (proteins). That is, having identified a particular differentially expressed gene as important in a particular state, screening of modulators of the expression of the gene or the gene product itself can be done. The gene products of differentially expressed genes are sometimes referred to herein as "prostate cancer proteins." The prostate cancer protein may be a fragment, or alternatively, be the full length protein to a fragment shown herein.

In one embodiment, screening for modulators of expression of specific genes is performed. Typically, the expression of only one or a few genes are evaluated. In another embodiment, screens are designed to first find compounds that bind to differentially expressed proteins. These compounds are then evaluated for the ability to modulate differentially expressed activity. Moreover, once initial candidate compounds are identified, variants can be further screened to better evaluate structure activity relationships.

In a preferred embodiment, binding assays are done. In general, purified or isolated gene product is used; that is, the gene products of one or more differentially expressed nucleic acids are made. For example, antibodies are generated to the protein gene products, and standard immunoassays are run to determine the amount of protein present. Alternatively, cells comprising the prostate cancer proteins can be used in the assays.

Thus, in a preferred embodiment, the methods comprise combining a prostate cancer protein and a candidate compound, and determining the binding of the compound to the prostate cancer protein. Preferred embodiments utilize the human prostate cancer protein,

although other mammalian proteins may also be used, e.g. for the development of animal models of human disease. In some embodiments, as outlined herein, variant or derivative prostate cancer proteins may be used.

Generally, in a preferred embodiment of the methods herein, the prostate cancer protein or the candidate agent is non-diffusably bound to an insoluble support having isolated sample receiving areas (e.g. a microtiter plate, an array, etc.). The insoluble supports may be made of any composition to which the compositions can be bound, is readily separated from soluble material, and is otherwise compatible with the overall method of screening. The surface of such supports may be solid or porous and of any convenient shape. Examples of suitable insoluble supports include microtiter plates, arrays, membranes and beads. These are typically made of glass, plastic (e.g., polystyrene), polysaccharides, nylon or nitrocellulose, teflon™, etc. Microtiter plates and arrays are especially convenient because a large number of assays can be carried out simultaneously, using small amounts of reagents and samples. The particular manner of binding of the composition is not crucial so long as it is compatible with the reagents and overall methods of the invention, maintains the activity of the composition and is nondiffusable. Preferred methods of binding include the use of antibodies (which do not sterically block either the ligand binding site or activation sequence when the protein is bound to the support), direct binding to "sticky" or ionic supports, chemical crosslinking, the synthesis of the protein or agent on the surface, etc. Following binding of the protein or agent, excess unbound material is removed by washing. The sample receiving areas may then be blocked through incubation with bovine serum albumin (BSA), casein or other innocuous protein or other moiety.

In a preferred embodiment, the prostate cancer protein is bound to the support, and a test compound is added to the assay. Alternatively, the candidate agent is bound to the support and the prostate cancer protein is added. Novel binding agents include specific antibodies, non-natural binding agents identified in screens of chemical libraries, peptide analogs, etc. Of particular interest are screening assays for agents that have a low toxicity for human cells. A wide variety of assays may be used for this purpose, including labeled in vitro protein-protein binding assays, electrophoretic mobility shift assays, immunoassays for protein binding, functional assays (phosphorylation assays, etc.) and the like.

The determination of the binding of the test modulating compound to the prostate cancer protein may be done in a number of ways. In a preferred embodiment, the compound is labeled, and binding determined directly, e.g., by attaching all or a portion of the prostate cancer protein to a solid support, adding a labeled candidate agent (e.g., a fluorescent label), washing off excess reagent, and determining whether the label is present on the solid support. Various blocking and washing steps may be utilized as appropriate.

In some embodiments, only one of the components is labeled, e.g., the proteins (or proteinaceous candidate compounds) can be labeled. Alternatively, more than one component can be labeled with different labels, e.g., ^{125}I for the proteins and a fluorophor for the compound. Proximity reagents, e.g., quenching or energy transfer reagents are also useful.

In one embodiment, the binding of the test compound is determined by competitive binding assay. The competitor is a binding moiety known to bind to the target molecule (i.e., a prostate cancer protein), such as an antibody, peptide, binding partner, ligand, etc. Under certain circumstances, there may be competitive binding between the compound and the binding moiety, with the binding moiety displacing the compound. In one embodiment, the test compound is labeled. Either the compound, or the competitor, or both, is added first to the protein for a time sufficient to allow binding, if present. Incubations may be performed at a temperature which facilitates optimal activity, typically between 4 and 40°C. Incubation periods are typically optimized, e.g., to facilitate rapid high throughput screening. Typically between 0.1 and 1 hour will be sufficient. Excess reagent is generally removed or washed away. The second component is then added, and the presence or absence of the labeled component is followed, to indicate binding.

In a preferred embodiment, the competitor is added first, followed by the test compound. Displacement of the competitor is an indication that the test compound is binding to the prostate cancer protein and thus is capable of binding to, and potentially modulating, the activity of the prostate cancer protein. In this embodiment, either component can be labeled. Thus, e.g., if the competitor is labeled, the presence of label in the wash solution indicates displacement by the agent. Alternatively, if the test compound is labeled, the presence of the label on the support indicates displacement.

In an alternative embodiment, the test compound is added first, with incubation and washing, followed by the competitor. The absence of binding by the competitor may indicate that the test compound is bound to the prostate cancer protein with a higher affinity. Thus, if the test compound is labeled, the presence of the label on the support, coupled with a lack of competitor binding, may indicate that the test compound is capable of binding to the prostate cancer protein.

In a preferred embodiment, the methods comprise differential screening to identity agents that are capable of modulating the activity of the prostate cancer proteins. In this embodiment, the methods comprise combining a prostate cancer protein and a competitor in a first sample. A second sample comprises a test compound, a prostate cancer protein, and a competitor. The binding of the competitor is determined for both samples, and a change, or difference in binding between the two samples indicates the presence of an agent capable of binding to the prostate cancer protein and potentially modulating its activity. That is, if the binding of the competitor is different in the second sample relative to the first sample, the agent is capable of binding to the prostate cancer protein.

Alternatively, differential screening is used to identify drug candidates that bind to the native prostate cancer protein, but cannot bind to modified prostate cancer proteins. The structure of the prostate cancer protein may be modeled, and used in rational drug design to synthesize agents that interact with that site. Drug candidates that affect the activity of a prostate cancer protein are also identified by screening drugs for the ability to either enhance or reduce the activity of the protein.

Positive controls and negative controls may be used in the assays. Preferably control and test samples are performed in at least triplicate to obtain statistically significant results. Incubation of all samples is for a time sufficient for the binding of the agent to the protein. Following incubation, samples are washed free of non-specifically bound material and the amount of bound, generally labeled agent determined. For example, where a radiolabel is employed, the samples may be counted in a scintillation counter to determine the amount of bound compound.

A variety of other reagents may be included in the screening assays. These include reagents like salts, neutral proteins, e.g. albumin, detergents, etc. which may be used to facilitate optimal protein-protein binding and/or reduce non-specific or background

interactions. Also reagents that otherwise improve the efficiency of the assay, such as protease inhibitors, nuclease inhibitors, anti-microbial agents, etc., may be used. The mixture of components may be added in an order that provides for the requisite binding.

In a preferred embodiment, the invention provides methods for screening for a compound capable of modulating the activity of a prostate cancer protein. The methods comprise adding a test compound, as defined above, to a cell comprising prostate cancer proteins. Preferred cell types include almost any cell. The cells contain a recombinant nucleic acid that encodes a prostate cancer protein. In a preferred embodiment, a library of candidate agents are tested on a plurality of cells.

In one aspect, the assays are evaluated in the presence or absence or previous or subsequent exposure of physiological signals, e.g. hormones, antibodies, peptides, antigens, cytokines, growth factors, action potentials, pharmacological agents including chemotherapeutics, radiation, carcinogenics, or other cells (i.e. cell-cell contacts). In another example, the determinations are determined at different stages of the cell cycle process.

In this way, compounds that modulate prostate cancer agents are identified. Compounds with pharmacological activity are able to enhance or interfere with the activity of the prostate cancer protein. Once identified, similar structures are evaluated to identify critical structural feature of the compound.

In one embodiment, a method of inhibiting prostate cancer cell division is provided. The method comprises administration of a prostate cancer inhibitor. In another embodiment, a method of inhibiting prostate cancer is provided. The method comprises administration of a prostate cancer inhibitor. In a further embodiment, methods of treating cells or individuals with prostate cancer are provided. The method comprises administration of a prostate cancer inhibitor.

In one embodiment, a prostate cancer inhibitor is an antibody as discussed above. In another embodiment, the prostate cancer inhibitor is an antisense molecule.

A variety of cell growth, proliferation, and metastasis assays are known to those of skill in the art, as described below.

Soft agar growth or colony formation in suspension

Normal cells require a solid substrate to attach and grow. When the cells are transformed, they lose this phenotype and grow detached from the substrate. For example,

transformed cells can grow in stirred suspension culture or suspended in semi-solid media, such as semi-solid or soft agar. The transformed cells, when transfected with tumor suppressor genes, regenerate normal phenotype and require a solid substrate to attach and grow. Soft agar growth or colony formation in suspension assays can be used to identify modulators of prostate cancer sequences, which when expressed in host cells, inhibit abnormal cellular proliferation and transformation. A therapeutic compound would reduce or eliminate the host cells' ability to grow in stirred suspension culture or suspended in semi-solid media, such as semi-solid or soft.

Techniques for soft agar growth or colony formation in suspension assays are described in Freshney, *Culture of Animal Cells a Manual of Basic Technique* (3rd ed., 1994), herein incorporated by reference. *See also*, the methods section of Garkavtsev *et al.* (1996), *supra*, herein incorporated by reference.

Contact inhibition and density limitation of growth

Normal cells typically grow in a flat and organized pattern in a petri dish until they touch other cells. When the cells touch one another, they are contact inhibited and stop growing. When cells are transformed, however, the cells are not contact inhibited and continue to grow to high densities in disorganized foci. Thus, the transformed cells grow to a higher saturation density than normal cells. This can be detected morphologically by the formation of a disoriented monolayer of cells or rounded cells in foci within the regular pattern of normal surrounding cells. Alternatively, labeling index with (^{3}H)-thymidine at saturation density can be used to measure density limitation of growth. *See* Freshney (1994), *supra*. The transformed cells, when transfected with tumor suppressor genes, regenerate a normal phenotype and become contact inhibited and would grow to a lower density.

In this assay, labeling index with (^{3}H)-thymidine at saturation density is a preferred method of measuring density limitation of growth. Transformed host cells are transfected with a prostate cancer-associated sequence and are grown for 24 hours at saturation density in non-limiting medium conditions. The percentage of cells labeling with (^{3}H)-thymidine is determined autoradiographically. *See*, Freshney (1994), *supra*.

Growth factor or serum dependence

Transformed cells have a lower serum dependence than their normal counterparts (*see, e.g.,* Temin, *J. Natl. Cancer Insti.* 37:167-175 (1966); Eagle *et al., J. Exp. Med.* 131:836-879 (1970)); Freshney, *supra.* This is in part due to release of various growth factors by the transformed cells. Growth factor or serum dependence of transformed host cells can be compared with that of control.

Tumor specific markers levels

Tumor cells release an increased amount of certain factors (hereinafter "tumor specific markers") than their normal counterparts. For example, plasminogen activator (PA) is released from human glioma at a higher level than from normal brain cells (*see, e.g.,* Gullino, *Angiogenesis, tumor vascularization, and potential interference with tumor growth.* in *Biological Responses in Cancer*, pp. 178-184 (Mihich (ed.) 1985)). Similarly, Tumor angiogenesis factor (TAF) is released at a higher level in tumor cells than their normal counterparts. *See, e.g.,* Folkman, *Angiogenesis and Cancer, Sem Cancer Biol.* (1992)).

Various techniques which measure the release of these factors are described in Freshney (1994), *supra.* Also, *see*, Unkless *et al.* , *J. Biol. Chem.* 249:4295-4305 (1974); Strickland & Beers, *J. Biol. Chem.* 251:5694-5702 (1976); Whur *et al., Br. J. Cancer* 42:305-312 (1980); Gullino, *Angiogenesis, tumor vascularization, and potential interference with tumor growth.* in *Biological Responses in Cancer*, pp. 178-184 (Mihich (ed.) 1985); Freshney *Anticancer Res.* 5:111-130 (1985).

Invasiveness into Matrigel

The degree of invasiveness into Matrigel or some other extracellular matrix constituent can be used as an assay to identify compounds that modulate prostate cancer-associated sequences. Tumor cells exhibit a good correlation between malignancy and invasiveness of cells into Matrigel or some other extracellular matrix constituent. In this assay, tumorigenic cells are typically used as host cells. Expression of a tumor suppressor gene in these host cells would decrease invasiveness of the host cells.

Techniques described in Freshney (1994), *supra*, can be used. Briefly, the level of invasion of host cells can be measured by using filters coated with Matrigel or some

other extracellular matrix constituent. Penetration into the gel, or through to the distal side of the filter, is rated as invasiveness, and rated histologically by number of cells and distance moved, or by prelabeling the cells with ^{125}I and counting the radioactivity on the distal side of the filter or bottom of the dish. *See, e.g.,* Freshney (1984), *supra.*

Tumor growth in vivo

Effects of prostate cancer-associated sequences on cell growth can be tested in transgenic or immune-suppressed mice. Knock-out transgenic mice can be made, in which the prostate cancer gene is disrupted or in which a prostate cancer gene is inserted. Knock-out transgenic mice can be made by insertion of a marker gene or other heterologous gene into the endogenous prostate cancer gene site in the mouse genome via homologous recombination. Such mice can also be made by substituting the endogenous prostate cancer gene with a mutated version of the prostate cancer gene, or by mutating the endogenous prostate cancer gene, e.g., by exposure to carcinogens.

A DNA construct is introduced into the nuclei of embryonic stem cells. Cells containing the newly engineered genetic lesion are injected into a host mouse embryo, which is re-implanted into a recipient female. Some of these embryos develop into chimeric mice that possess germ cells partially derived from the mutant cell line. Therefore, by breeding the chimeric mice it is possible to obtain a new line of mice containing the introduced genetic lesion (*see, e.g.,* Capecchi *et al., Science* 244:1288 (1989)). Chimeric targeted mice can be derived according to Hogan *et al., Manipulating the Mouse Embryo: A Laboratory Manual,* Cold Spring Harbor Laboratory (1988) and *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach,* Robertson, ed., IRL Press, Washington, D.C., (1987).

Alternatively, various immune-suppressed or immune-deficient host animals can be used. For example, genetically athymic "nude" mouse (*see, e.g.,* Giovanella *et al., J. Natl. Cancer Inst.* 52:921 (1974)), a SCID mouse, a thymectomized mouse, or an irradiated mouse (*see, e.g.,* Bradley *et al., Br. J. Cancer* 38:263 (1978); Selby *et al., Br. J. Cancer* 41:52 (1980)) can be used as a host. Transplantable tumor cells (typically about 10^6 cells) injected into isogenic hosts will produce invasive tumors in a high proportions of cases, while normal cells of similar origin will not. In hosts which developed invasive tumors, cells expressing a prostate cancer-associated sequences are injected subcutaneously. After a

suitable length of time, preferably 4-8 weeks, tumor growth is measured (e.g., by volume or by its two largest dimensions) and compared to the control. Tumors that have statistically significant reduction (using, e.g., Student's T test) are said to have inhibited growth.

Methods of identifying variant prostate cancer-associated sequences

Without being bound by theory, expression of various prostate cancer sequences is correlated with prostate cancer. Accordingly, disorders based on mutant or variant prostate cancer genes may be determined. In one embodiment, the invention provides methods for identifying cells containing variant prostate cancer genes, e.g., determining all or part of the sequence of at least one endogenous prostate cancer genes in a cell. This may be accomplished using any number of sequencing techniques. In a preferred embodiment, the invention provides methods of identifying the prostate cancer genotype of an individual, e.g., determining all or part of the sequence of at least one prostate cancer gene of the individual. This is generally done in at least one tissue of the individual, and may include the evaluation of a number of tissues or different samples of the same tissue. The method may include comparing the sequence of the sequenced prostate cancer gene to a known prostate cancer gene, i.e., a wild-type gene.

The sequence of all or part of the prostate cancer gene can then be compared to the sequence of a known prostate cancer gene to determine if any differences exist. This can be done using any number of known homology programs, such as Bestfit, etc. In a preferred embodiment, the presence of a difference in the sequence between the prostate cancer gene of the patient and the known prostate cancer gene correlates with a disease state or a propensity for a disease state, as outlined herein.

In a preferred embodiment, the prostate cancer genes are used as probes to determine the number of copies of the prostate cancer gene in the genome.

In another preferred embodiment, the prostate cancer genes are used as probes to determine the chromosomal localization of the prostate cancer genes. Information such as chromosomal localization finds use in providing a diagnosis or prognosis in particular when chromosomal abnormalities such as translocations, and the like are identified in the prostate cancer gene locus.

Administration of pharmaceutical and vaccine compositions

In one embodiment, a therapeutically effective dose of a prostate cancer protein or modulator thereof, is administered to a patient. By "therapeutically effective dose" herein is meant a dose that produces effects for which it is administered. The exact dose will depend on the purpose of the treatment, and will be ascertainable by one skilled in the art using known techniques (e.g., Ansel *et al.*, *Pharmaceutical Dosage Forms and Drug Delivery*; Lieberman, *Pharmaceutical Dosage Forms* (vols. 1-3, 1992), Dekker, ISBN 0824770846, 082476918X, 0824712692, 0824716981; Lloyd, *The Art, Science and Technology of Pharmaceutical Compounding* (1999); and Pickar, *Dosage Calculations* (1999)). As is known in the art, adjustments for prostate cancer degradation, systemic versus localized delivery, and rate of new protease synthesis, as well as the age, body weight, general health, sex, diet, time of administration, drug interaction and the severity of the condition may be necessary, and will be ascertainable with routine experimentation by those skilled in the art. U.S. Patent Application N. 09/687,576, further discloses the use of compositions and methods of diagnosis and treatment in prostate cancer is hereby expressly incorporated by reference.

A "patient" for the purposes of the present invention includes both humans and other animals, particularly mammals. Thus the methods are applicable to both human therapy and veterinary applications. In the preferred embodiment the patient is a mammal, preferably a primate, and in the most preferred embodiment the patient is human.

The administration of the prostate cancer proteins and modulators thereof of the present invention can be done in a variety of ways as discussed above, including, but not limited to, orally, subcutaneously, intravenously, intranasally, transdermally, intraperitoneally, intramuscularly, intrapulmonary, vaginally, rectally, or intraocularly. In some instances, e.g., in the treatment of wounds and inflammation, the prostate cancer proteins and modulators may be directly applied as a solution or spray.

The pharmaceutical compositions of the present invention comprise a prostate cancer protein in a form suitable for administration to a patient. In the preferred embodiment, the pharmaceutical compositions are in a water soluble form, such as being present as pharmaceutically acceptable salts, which is meant to include both acid and base addition salts. "Pharmaceutically acceptable acid addition salt" refers to those salts that retain the

biological effectiveness of the free bases and that are not biologically or otherwise undesirable, formed with inorganic acids such as hydrochloric acid, hydrobromic acid, sulfuric acid, nitric acid, phosphoric acid and the like, and organic acids such as acetic acid, propionic acid, glycolic acid, pyruvic acid, oxalic acid, maleic acid, malonic acid, succinic acid, fumaric acid, tartaric acid, citric acid, benzoic acid, cinnamic acid, mandelic acid, methanesulfonic acid, ethanesulfonic acid, p-toluenesulfonic acid, salicylic acid and the like. "Pharmaceutically acceptable base addition salts" include those derived from inorganic bases such as sodium, potassium, lithium, ammonium, calcium, magnesium, iron, zinc, copper, manganese, aluminum salts and the like. Particularly preferred are the ammonium, potassium, sodium, calcium, and magnesium salts. Salts derived from pharmaceutically acceptable organic non-toxic bases include salts of primary, secondary, and tertiary amines, substituted amines including naturally occurring substituted amines, cyclic amines and basic ion exchange resins, such as isopropylamine, trimethylamine, diethylamine, triethylamine, tripropylamine, and ethanolamine.

The pharmaceutical compositions may also include one or more of the following: carrier proteins such as serum albumin; buffers; fillers such as microcrystalline cellulose, lactose, corn and other starches; binding agents; sweeteners and other flavoring agents; coloring agents; and polyethylene glycol.

The pharmaceutical compositions can be administered in a variety of unit dosage forms depending upon the method of administration. For example, unit dosage forms suitable for oral administration include, but are not limited to, powder, tablets, pills, capsules and lozenges. It is recognized that prostate cancer protein modulators (e.g., antibodies, antisense constructs, ribozymes, small organic molecules, *etc.*) when administered orally, should be protected from digestion. This is typically accomplished either by complexing the molecule(s) with a composition to render it resistant to acidic and enzymatic hydrolysis, or by packaging the molecule(s) in an appropriately resistant carrier, such as a liposome or a protection barrier. Means of protecting agents from digestion are well known in the art.

The compositions for administration will commonly comprise a prostate cancer protein modulator dissolved in a pharmaceutically acceptable carrier, preferably an aqueous carrier. A variety of aqueous carriers can be used, e.g., buffered saline and the like. These solutions are sterile and generally free of undesirable matter. These compositions may

be sterilized by conventional, well known sterilization techniques. The compositions may contain pharmaceutically acceptable auxiliary substances as required to approximate physiological conditions such as pH adjusting and buffering agents, toxicity adjusting agents and the like, e.g., sodium acetate, sodium chloride, potassium chloride, calcium chloride, sodium lactate and the like. The concentration of active agent in these formulations can vary widely, and will be selected primarily based on fluid volumes, viscosities, body weight and the like in accordance with the particular mode of administration selected and the patient's needs (e.g., *Remington's Pharmaceutical Science* (15th ed., 1980) and Goodman & Gillman, *The Pharmacologial Basis of Therapeutics* (Hardman *et al.*,eds., 1996)).

Thus, a typical pharmaceutical composition for intravenous administration would be about 0.1 to 10 mg per patient per day. Dosages from 0.1 up to about 100 mg per patient per day may be used, particularly when the drug is administered to a secluded site and not into the blood stream, such as into a body cavity or into a lumen of an organ. Substantially higher dosages are possible in topical administration. Actual methods for preparing parenterally administrable compositions will be known or apparent to those skilled in the art, *e.g.*, *Remington's Pharmaceutical Science* and Goodman and Gillman, *The Pharmacologial Basis of Therapeutics*, *supra*.

The compositions containing modulators of prostate cancer proteins can be administered for therapeutic or prophylactic treatments. In therapeutic applications, compositions are administered to a patient suffering from a disease (e.g., a cancer) in an amount sufficient to cure or at least partially arrest the disease and its complications. An amount adequate to accomplish this is defined as a "therapeutically effective dose." Amounts effective for this use will depend upon the severity of the disease and the general state of the patient's health. Single or multiple administrations of the compositions may be administered depending on the dosage and frequency as required and tolerated by the patient. In any event, the composition should provide a sufficient quantity of the agents of this invention to effectively treat the patient. An amount of modulator that is capable of preventing or slowing the development of cancer in a mammal is referred to as a "prophylactically effective dose." The particular dose required for a prophylactic treatment will depend upon the medical condition and history of the mammal, the particular cancer being prevented, as well as other factors such as age, weight, gender, administration route, efficiency, *etc.* Such prophylactic

treatments may be used, *e.g.*, in a mammal who has previously had cancer to prevent a recurrence of the cancer, or in a mammal who is suspected of having a significant likelihood of developing cancer.

It will be appreciated that the present prostate cancer protein-modulating compounds can be administered alone or in combination with additional prostate cancer modulating compounds or with other therapeutic agent, *e.g.*, other anti-cancer agents or treatments.

In numerous embodiments, one or more nucleic acids, *e.g.*, polynucleotides comprising nucleic acid sequences set forth in Tables 1-16, such as antisense polynucleotides or ribozymes, will be introduced into cells, *in vitro* or *in vivo*. The present invention provides methods, reagents, vectors, and cells useful for expression of prostate cancer-associated polypeptides and nucleic acids using *in vitro* (cell-free), *ex vivo* or *in vivo* (cell or organism-based) recombinant expression systems.

The particular procedure used to introduce the nucleic acids into a host cell for expression of a protein or nucleic acid is application specific. Many procedures for introducing foreign nucleotide sequences into host cells may be used. These include the use of calcium phosphate transfection, spheroplasts, electroporation, liposomes, microinjection, plasma vectors, viral vectors and any of the other well known methods for introducing cloned genomic DNA, cDNA, synthetic DNA or other foreign genetic material into a host cell (*see, e.g.*, Berger & Kimmel, *Guide to Molecular Cloning Techniques, Methods in Enzymology* volume 152 (Berger), Ausubel *et al.*, eds., *Current Protocols* (supplemented through 1999), and Sambrook *et al.*, *Molecular Cloning - A Laboratory Manual* (2nd ed., Vol. 1-3, 1989.

In a preferred embodiment, prostate cancer proteins and modulators are administered as therapeutic agents, and can be formulated as outlined above. Similarly, prostate cancer genes (including both the full-length sequence, partial sequences, or regulatory sequences of the prostate cancer coding regions) can be administered in a gene therapy application. These prostate cancer genes can include antisense applications, either as gene therapy (i.e. for incorporation into the genome) or as antisense compositions, as will be appreciated by those in the art.

Prostate cancer polypeptides and polynucleotides can also be administered as vaccine compositions to stimulate HTL, CTL and antibody responses.. Such vaccine

compositions can include, e.g., lipidated peptides (*see, e.g.*,Vitiello, A. *et al.*, *J. Clin. Invest.* 95:341 (1995)), peptide compositions encapsulated in poly(DL-lactide-co-glycolide) ("PLG") microspheres (*see, e.g.*, Eldridge, *et al.*, *Molec. Immunol.* 28:287-294, (1991); Alonso *et al.*, *Vaccine* 12:299-306 (1994); Jones *et al.*, *Vaccine* 13:675-681 (1995)), peptide compositions contained in immune stimulating complexes (ISCOMS) (*see, e.g.*, Takahashi *et al.*, *Nature* 344:873-875 (1990); Hu *et al.*, *Clin Exp Immunol.* 113:235-243 (1998)), multiple antigen peptide systems (MAPs) (*see, e.g.*, Tam, *Proc. Natl. Acad. Sci. U.S.A.* 85:5409-5413 (1988); Tam, *J. Immunol. Methods* 196:17-32 (1996)), peptides formulated as multivalent peptides; peptides for use in ballistic delivery systems, typically crystallized peptides, viral delivery vectors (Perkus, *et al.*, In: *Concepts in vaccine development* (Kaufmann, ed., p. 379, 1996); Chakrabarti, *et al.*, *Nature* 320:535 (1986); Hu *et al.*, *Nature* 320:537 (1986); Kieny, *et al.*, *AIDS Bio/Technology* 4:790 (1986); Top *et al.*, *J. Infect. Dis.* 124:148 (1971); Chanda *et al.*, *Virology* 175:535 (1990)), particles of viral or synthetic origin (*see, e.g.*, Kofler *et al.*, *J. Immunol. Methods.* 192:25 (1996); Eldridge *et al.*, *Sem. Hematol.* 30:16 (1993); Falo *et al.*, *Nature Med.* 7:649 (1995)), adjuvants (Warren *et al.*, *Annu. Rev. Immunol.* 4:369 (1986); Gupta *et al.*, *Vaccine* 11:293 (1993)), liposomes (Reddy *et al.*, *J. Immunol.* 148:1585 (1992); Rock, *Immunol. Today* 17:131 (1996)), or, naked or particle absorbed cDNA (Ulmer, *et al.*, *Science* 259:1745 (1993); Robinson *et al.*, *Vaccine* 11:957 (1993); Shiver *et al.*, In: *Concepts in vaccine development* (Kaufmann, ed., p. 423, 1996); Cease & Berzofsky, *Annu. Rev. Immunol.* 12:923 (1994) and Eldridge *et al.*, *Sem. Hematol.* 30:16 (1993)). Toxin-targeted delivery technologies, also known as receptor mediated targeting, such as those of Avant Immunotherapeutics, Inc. (Needham, Massachusetts) may also be used.

Vaccine compositions often include adjuvants. Many adjuvants contain a substance designed to protect the antigen from rapid catabolism, such as aluminum hydroxide or mineral oil, and a stimulator of immune responses, such as lipid A, *Bortadella pertussis* or *Mycobacterium tuberculosis* derived proteins. Certain adjuvants are commercially available as, e.g., Freund's Incomplete Adjuvant and Complete Adjuvant (Difco Laboratories, Detroit, MI); Merck Adjuvant 65 (Merck and Company, Inc., Rahway, NJ); AS-2 (SmithKline Beecham, Philadelphia, PA); aluminum salts such as aluminum hydroxide gel (alum) or aluminum phosphate; salts of calcium, iron or zinc; an insoluble suspension of acylated tyrosine; acylated sugars; cationically or anionically derivatized polysaccharides;

polyphosphazenes; biodegradable microspheres; monophosphoryl lipid A and quil A. Cytokines, such as GM-CSF, interleukin-2, -7, -12, and other like growth factors, may also be used as adjuvants.

Vaccines can be administered as nucleic acid compositions wherein DNA or RNA encoding one or more of the polypeptides, or a fragment thereof, is administered to a patient. This approach is described, for instance, in Wolff *et. al.*, *Science* 247:1465 (1990) as well as U.S. Patent Nos. 5,580,859; 5,589,466; 5,804,566; 5,739,118; 5,736,524; 5,679,647; WO 98/04720; and in more detail below. Examples of DNA-based delivery technologies include "naked DNA", facilitated (bupivicaine, polymers, peptide-mediated) delivery, cationic lipid complexes, and particle-mediated ("gene gun") or pressure-mediated delivery (*see, e.g.,* U.S. Patent No. 5,922,687).

For therapeutic or prophylactic immunization purposes, the peptides of the invention can be expressed by viral or bacterial vectors. Examples of expression vectors include attenuated viral hosts, such as vaccinia or fowlpox. This approach involves the use of vaccinia virus, e.g., as a vector to express nucleotide sequences that encode prostate cancer polypeptides or polypeptide fragments. Upon introduction into a host, the recombinant vaccinia virus expresses the immunogenic peptide, and thereby elicits an immune response. Vaccinia vectors and methods useful in immunization protocols are described in, *e.g.*, U.S. Patent No. 4,722,848. Another vector is BCG (Bacille Calmette Guerin). BCG vectors are described in Stover *et al.*, *Nature* 351:456-460 (1991). A wide variety of other vectors useful for therapeutic administration or immunization e.g. adeno and adeno-associated virus vectors, retroviral vectors, *Salmonella typhi* vectors, detoxified anthrax toxin vectors, and the like, will be apparent to those skilled in the art from the description herein (*see, e.g.,* Shata *et al.*, *Mol Med Today* 6:66-71 (2000); Shedlock *et al.*, *J Leukoc Biol* 68:793-806 (2000); Hipp *et al.*, *In Vivo* 14:571-85 (2000)).

Methods for the use of genes as DNA vaccines are well known, and include placing a prostate cancer gene or portion of a prostate cancer gene under the control of a regulatable promoter or a tissue-specific promoter for expression in a prostate cancer patient. The prostate cancer gene used for DNA vaccines can encode full-length prostate cancer proteins, but more preferably encodes portions of the prostate cancer proteins including peptides derived from the prostate cancer protein. In one embodiment, a patient is

immunized with a DNA vaccine comprising a plurality of nucleotide sequences derived from a prostate cancer gene. For example, prostate cancer-associated genes or sequence encoding subfragments of a prostate cancer protein are introduced into expression vectors and tested for their immunogenicity in the context of Class I MHC and an ability to generate cytotoxic T cell responses. This procedure provides for production of cytotoxic T cell responses against cells which present antigen, including intracellular epitopes.

In a preferred embodiment, the DNA vaccines include a gene encoding an adjuvant molecule with the DNA vaccine. Such adjuvant molecules include cytokines that increase the immunogenic response to the prostate cancer polypeptide encoded by the DNA vaccine. Additional or alternative adjuvants are available.

In another preferred embodiment prostate cancer genes find use in generating animal models of prostate cancer. When the prostate cancer gene identified is repressed or diminished in cancer tissue, gene therapy technology, e.g., wherein antisense RNA directed to the prostate cancer gene will also diminish or repress expression of the gene. Animal models of prostate cancer find use in screening for modulators of a prostate cancer-associated sequence or modulators of prostate cancer. Similarly, transgenic animal technology including gene knockout technology, e.g. as a result of homologous recombination with an appropriate gene targeting vector, will result in the absence or increased expression of the prostate cancer protein. When desired, tissue-specific expression or knockout of the prostate cancer protein may be necessary.

It is also possible that the prostate cancer protein is overexpressed in prostate cancer. As such, transgenic animals can be generated that overexpress the prostate cancer protein. Depending on the desired expression level, promoters of various strengths can be employed to express the transgene. Also, the number of copies of the integrated transgene can be determined and compared for a determination of the expression level of the transgene. Animals generated by such methods find use as animal models of prostate cancer and are additionally useful in screening for modulators to treat prostate cancer.

Kits for Use in Diagnostic and/or Prognostic Applications

For use in diagnostic, research, and therapeutic applications suggested above, kits are also provided by the invention. In the diagnostic and research applications such kits

may include any or all of the following: assay reagents, buffers, prostate cancer-specific nucleic acids or antibodies, hybridization probes and/or primers, antisense polynucleotides, ribozymes, dominant negative prostate cancer polypeptides or polynucleotides, small molecules inhibitors of prostate cancer-associated sequences *etc.* A therapeutic product may include sterile saline or another pharmaceutically acceptable emulsion and suspension base.

In addition, the kits may include instructional materials containing directions (i.e., protocols) for the practice of the methods of this invention. While the instructional materials typically comprise written or printed materials they are not limited to such. Any medium capable of storing such instructions and communicating them to an end user is contemplated by this invention. Such media include, but are not limited to electronic storage media (e.g., magnetic discs, tapes, cartridges, chips), optical media (e.g., CD ROM), and the like. Such media may include addresses to internet sites that provide such instructional materials.

The present invention also provides for kits for screening for modulators of prostate cancer-associated sequences. Such kits can be prepared from readily available materials and reagents. For example, such kits can comprise one or more of the following materials: a prostate cancer-associated polypeptide or polynucleotide, reaction tubes, and instructions for testing prostate cancer-associated activity. Optionally, the kit contains biologically active prostate cancer protein. A wide variety of kits and components can be prepared according to the present invention, depending upon the intended user of the kit and the particular needs of the user. Diagnosis would typically involve evaluation of a plurality of genes or products. The genes will be selected based on correlations with important parameters in disease which may be identified in historical or outcome data.

EXAMPLES

Example 1: Tissue Preparation, Labeling Chips, and Fingerprints

Purifying total RNA from tissue sample using TRIzol Reagent

The sample weight is first estimated. The tissue samples are homogenized in 1 ml of TRIzol per 50 mg of tissue using a homogenizer (e.g., Polytron 3100). The size of the generator/probe used depends upon the sample amount. A generator that is too large for the amount of tissue to be homogenized will cause a loss of sample and lower RNA yield. A larger generator (e.g., 20 mm) is suitable for tissue samples weighing more than 0.6 g. Fill tubes should not be overfilled. If the working volume is greater than 2 ml and no greater than 10 ml, a 15 ml polypropylene tube (Falcon 2059) is suitable for homogenization.

Tissues should be kept frozen until homogenized. The TRIzol is added directly to the frozen tissue before homogenization. Following homogenization, the insoluble material is removed from the homogenate by centrifugation at 7500 x g for 15 min. in a Sorvall superspeed or 12,000 x g for 10 min. in an Eppendorf centrifuge at 4°C. The cleared homogenate is then transferred to a new tube(s). Samples may be frozen and stored at -60 to -70°C for at least one month or else continue with the purification.

The next process is phase separation. The homogenized samples are incubated for 5 minutes at room temperature. Then, 0.2 ml of chloroform per 1ml of TRIzol reagent is added to the homogenization mixture. The tubes are securely capped and shaken vigorously by hand (do not vortex) for 15 seconds. The samples are then incubated at room temp. for 2-3 minutes and next centrifuged at 6500 rpm in a Sorvall superspeed for 30 min. at 4oC.

The next process is RNA Precipitation. The aqueous phase is transferred to a fresh tube. The organic phase can be saved if isolation of DNA or protein is desired. Then 0.5 ml of isopropyl alcohol is added per 1ml of TRIzol reagent used in the original homogenization. Then, the tubes are securely capped and inverted to mix. The samples are then incubated at room temp. for 10 minutes an centrifuged at 6500 rpm in Sorvall for 20 min. at 4°C.

The RNA is then washed. The supernatant is poured off and the pellet washed with cold 75% ethanol. 1 ml of 75% ethanol is used per 1 ml of the TRIzol reagent used in the initial homogenization. The tubes are capped securely and inverted several times to loosen pellet without vortexing . They are next centrifuged at <8000 rpm (<7500 x g) for 5 minutes at 4°C.

The RNA wash is decanted. The pellet is carefully transferred to an Eppendorf tube (sliding down the tube into the new tube by use of a pipet tip to help guide it in if necessary). Tube(s) sizes for precipitating the RNA depending on the working volumes. Larger tubes may take too long to dry. Dry pellet. The RNA is then resuspended in an appropriate volume (e.g., 2 -5 ug/ul) of DEPC H_20. The absorbance is then measured.

The poly A+ mRNA may next be purified from total RNA by other methods such as Qiagen' s RNeasy kit. The poly A^+ mRNA is purified from total RNA by adding the oligotex suspension which has been heated to 37°C and mixing prior to adding to RNA. The Elution Buffer is incubated at 70°C. If there is precipitate in the buffer, warm up the 2 x Binding Buffer at 65°C. The the total RNA is mixed with DEPC-treated water, 2 x Binding Buffer, and Oligotex according to Table 2 on page 16 of the Oligotex Handbook and next incubated for 3 minutes at 65°C and 10 minutes at room temperature.

The preparation is centrifuged for 2 minutes at 14,000 to 18,000 g, preferably, at a "soft setting," The supernatant is removed without disturbing Oligotex pellet. A little bit of solution can be left behind to reduce the loss of Oligotex. The supernatant is saved until satisfactory binding and elution of poly A^+ mRNA has been found.

Then, the preparation is gently resuspended in Wash Buffer OW2 and pipetted onto the spin column and centrifuged at full speed (soft setting if possible) for 1 minute.

Next, the spin column is transferred to a new collection tube and gently resuspended in Wash Buffer OW2 and centrifuged as described herein.

Then, the spin column is transferred to a new tube and eluted with 20 to 100 ul of preheated (70°C) Elution Buffer. The Oligotex resin is gently resuspended by pipetting up and down. The centrifugation is repeated as above and the elution repeated with fresh elution buffer or first eluate to keep the elution volume low.

The absorbance is next read to determine the yield, using diluted Elution Buffer as the blank.

Before proceeding with cDNA synthesis, the mRNA is precipitated before proceeding with cDNA synthesis, as components leftover or in the Elution Buffer from the Oligotex purification procedure will inhibit downstream enzymatic reactions of the mRNA. 0.4 vol. of 7.5 M NH4OAc + 2.5 vol. of cold 100% ethanol is added and the preparation precipitated at -20°C 1 hour to overnight (or 20-30 min. at -70°C), and centrifuged at 14,000-16,000 x g for 30 minutes at 4°C. Next, the pellet is washed with 0.5 ml of 80% ethanol (-20°C) and then centrifuged at 14,000-16,000 x g for 5 minutes at room temperature. The80% ethanol wash is then repeated. The last bit of ethanol from the pellet is then dried without use of a speed vacuum and the pellet is then resuspended in DEPC H_20 at 1ug/ul concentration.

Alternatively the RNA may be purified using other methods (e.g., Qiagen's RNeasy kit).

No more than 100 ug is added to the RNeasy column. The sample volume is adjusted to 100 ul with RNase-free water. 350 ul Buffer RLT and then 250 ul ethanol (100%) are added to the sample. The preparation is then mixed by pipetting and applied to an RNeasy mini spin column for centrifugation (15 sec at >10,000 rpm). If yield is low, reapply the flowthrough to the column and centrifuge again.

Then, transfer column to a new 2 ml collection tube and add 500 ul Buffer RPE and centrifuge for 15 sec at >10,000 rpm. The flowthrough is discarded. 500 ul Buffer RPE and is then added and the preparation is centriuged for 15 sec at >10,000 rpm. The flowthrough is discarded. and the column membrane dried by centrifuging for 2 min at maximum speed. The column is transferred to a new 1.5-ml collection tube. 30-50 ul of RNase-free water is applied directly onto column membrane. The column is then centrifuged for 1 min at >10,000 rpm and the elution step repeated.

The absorbance is then read to determine yield. If necessary, the material may be ethanol precipitated with ammonium acetate and 2.5X volume 100% ethanol.

First Strand cDNA Synthesis

The first strand can be make using using Gibco's "SuperScript Choice System for cDNA Synthesis" kit. The starting material is 5 ug of total RNA or 1 ug of polyA+ mRNAl. For total RNA, 2 ul of SuperScript RT is used; for polyA+ mRNA, 1 ul of SuperScript RT is used. The final volume of first strand synthesis mix is 20 ul. The RNA should be in a volume no greater than 10 ul. The RNA is incubated with 1 ul of 100 pmol T7-T24 oligo for 10 min at 70°C followed by addition on ice of 7 ul of: 4ul 5X 1st Strand Buffer, 2 ul of 0.1M DTT, and 1 ul of 10mM dNTP mix. The preparation is then incubated at 37°C for 2 min before addition of the SuperScript RT followed by incubation at 37°C for 1 hour.

Second Strand Synthesis

For the second strand synthesis, place 1st strand reactions on ice and add: 91 ul DEPC H_20; 30 ul 5X 2nd Strand Buffer; 3 ul 10mM dNTP mix; 1 ul 10 U/ul E.coli DNA Ligase; 4 ul 10 U/ul E.coli DNA Polymerase; and 1 ul 2 U/ul RNase H. Mix and incubate 2 hours at 16°C. Add 2 ul T4 DNA Polymerase. Incubate 5 min at 16°C. Add 10 ul of 0.5M EDTA.

Cleaning up cDNA

The cDNA is purified using Phenol:Chloroform:Isoamyl Alcohol (25:24:1) and Phase-Lock gel tubes. The PLG tubes are centrifuged for 30 sec at maximum speed. The cDNA mix is then transferred to PLG tube. An equal volume of phenol:chloroform:isamyl alcohol is then added, the preparation shaken vigorously (no vortexing), and centrifuged for 5 minutes at maximum speed. The top aqueous solution is transferred to a new tube and ethanol precipitated by adding 7.5X 5M NH4OAc and 2.5X volume of 100% ethanol. Next, it is centrifuged immediately at room temperature for 20 min, maximum speed. The supernatant is removed, and the pellet washed with 2X with cold 80% ethanol. As much ethanol wash as possible should be removed before air drying the pellet; and resuspending it in 3 ul RNase-free water.

In vitro Transcription (IVT) and labeling with biotin

In vitro Transcription (IVT) and labeling with biotin is performed as follows: Pipet 1.5 ul of cDNA into a thin-wall PCR tube. Make NTP labeling mix by combining 2 ul T7 10xATP (75 mM) (Ambion); 2 ul T7 10xGTP (75 mM) (Ambion); 1.5 ul T7 10xCTP (75 mM) (Ambion); 1.5 ul T7 10xUTP (75 mM) (Ambion); 3.75 ul 10 mM Bio-11-UTP (Boehringer-Mannheim/Roche or Enzo); 3.75 ul 10 mM Bio-16-CTP (Enzo); 2 ul 10x T7 transcription buffer (Ambion); and 2 ul 10x T7 enzyme mix (Ambion). The final volume is 20 ul. Incubate 6 hours at 37°C in a PCR machine. The RNA can be furthered cleaned. Clean-up follows the previous instructions for RNeasy columns or Qiagen's RNeasy protocol handbook. The cRNA often needs to be ethanol precipitated by resuspension in a volume compatible with the fragmentation step.

Fragmentation is performed as follows. 15 ug of labeled RNA is usually fragmented. Try to minimize the fragmentation reaction volume; a 10 ul volume is recommended but 20 ul is all right. Do not go higher than 20 ul because the magnesium in the fragmentation buffer contributes to precipitation in the hybridization buffer. Fragment RNA by incubation at 94 C for 35 minutes in 1 x Fragmentation buffer (5 x Fragmentation buffer is 200 mM Tris-acetate, pH 8.1; 500 mM KOAc; 150 mM MgOAc). The labeled RNA transcript can be analyzed before and after fragmentation. Samples can be heated to 65°C for 15 minutes and electrophoresed on 1% agarose/TBE gels to get an approximate idea of the transcript size range.

For hybridization, 200 ul (10 ug cRNA) of a hybridization mix is put on the chip. If multiple hybridizations are to be done (such as cycling through a 5 chip set), then it is recommended that an initial hybridization mix of 300 ul or more be made. The hybridization mix is: fragment labeled RNA (50 ng/ul final conc.); 50 pM 948-b control oligo; 1.5 pM BioB; 5 pM BioC; 25 pM BioD; 100 pM CRE; 0.1 mg/ml herring sperm DNA; 0.5 mg/ml acetylated BSA; and 300 ul with 1xMES hyb buffer.

The hybridization reaction is conducted with non-biotinylated IVT (purified by RNeasy columns) (see example 1 for steps from tissue to IVT): The following mixture is prepared:

IVT antisense RNA; 4 μg:	μl
Random Hexamers (1 μg/μl):	4 μl
H_2O:	μl
	14 μl

Incubate the above 14 μl mixture at 70°C for 10 min.; then put on ice.

The Reverse transcription procedure uses the following mixture:

0.1 M DTT:	3 μl
50X dNTP mix:	0.6 μl
H_2O:	2.4 μl
Cy3 or Cy5 dUTP (1mM):	3 μl
SS RT II (BRL):	1 μl
	16 μl

The above solution is added to the hybridization reaction and incubated for 30 min., 42°C. Then, 1 μl SSII is added and incubated for another hour before being placed on ice.

The 50X dNTP mix contains 25mM of cold dATP, dCTP, and dGTP, 10mM of dTTP and is made by adding 25 μl each of 100mM dATP, dCTP, and dGTP; 10 μl of 100mM dTTP to 15 μl H_2O.]

RNA degradation is performed as follows. Add 86 μl H2O, 1.5 μl 1M NaOH/ 2 mM EDTA and incubate at 65°C, 10 min.. For U-Con 30, 500 μl TE/sample spin at 7000 g for 10 min, save flow through for purification. For Qiagen purification, suspend u-con recovered material in 500 μl buffer PB and proceed using Qiagen protocol. For DNAse digestion, add 1 ul of 1/100 dilution of DNAse/30 ul Rx and incubate at 37°C for 15 min. Incubate at 5 min 95°C to denature the DNAse.

Sample preparation

For sample preparation, add Cot-1 DNA, 10 μl; 50X dNTPs, 1 μl; 20X SSC, 2.3 μl; Na pyro phosphate, 7.5 μl; 10 mg/ml Herring sperm DNA; 1 ul of 1/10 dilution to 21.8 final vol. Dry in speed vac. Resuspend in 15 μl H_20. Add 0.38 μl 10% SDS. Heat

95°C, 2 min and slow cool at room temp. for 20 min. Put on slide and hybridize overnight at 64°C. Washing after the hybridization: 3X SSC/0.03% SDS: 2 min., 37.5 ml 20X SSC+0.75ml 10% SDS in 250ml H_2O; 1X SSC: 5 min., 12.5 ml 20X SSC in 250ml H_2O; 0.2X SSC: 5 min., 2.5 ml 20X SSC in 250ml H_2O. Dry slides and scan at appropiate PMT's and channels.

Example 2: Taxol resistant Xenograft Model of Human Prostate Cancer

Treatment regimens that include paclitaxel (Taxol; Bristol-Myers Squibb Company, Princeton, NJ) have been particularly successful in treating hormone-refractory prostate cancer in the phase II setting (Smith et al., Semin. Oncol. 26(1 Suppl 2):109-11 (1999)). However, many patients develop tumors which are initially, or later become, resistant to taxol. To identify genes that may be involved with resistance to taxol, or are regulated in response to taxol resistance, and therefore may be used to treat, or identify, taxol resistance in patients, the following experiments were carried out.

The androgen-independent human cell line CWR22R was grown as a xenograft in nude mice (Nagabhushan et al., Cancer Res. 56(13):3042-3046 (1996); Agus et al., J. Natl. Cancer Inst.91(21):1869-1876 (1999); Bubendorf et al., J. Natl. Cancer Inst. 91(20):1758-1764 (1999)). Initially, these xenograft tumors were sensitive to therapeutic doses of taxol. The mice were treated continuously with sub-therapeutic doses, and the tumors were allowed to grow for 3-4 weeks, before surgical removal of the tumors. The tumor from an individual mouse was then minced, and a small portion was then injected into a healthy nude mouse, establishing a second

passage of the tumor. This mouse was then treated continuously with the same sub-therapeutic dose of taxol. This process was repeated 14 times, and a portion of each generation of xenograft tumor was collected. There was increasing resistance to therapeutic doses of taxol with each generation. Bythe end of the process, the tumors were fully resistant to therapeutic doses of taxol. RNA from each generation of tumor was then isolated, and individual mRNA species were quantified using a custom Affymetrix GeneChip® oligonucleotide microarray, with probes to interrogate approximately 35,000

unique mRNA transcripts. Genes were selected that showed a statistically significant up-regulation, or down-regulation, during the subsequent generations of increasingly taxol-resistant tumors. Only one gene was significantly up-regulated, whereas 24 genes were down-regulated; these are presented in Table 10.

The gene sequences identified to be overexpressed in prostate cancer may be used to identify coding regions from the public DNA database. The sequences may be used to either identify genes that encode known proteins, or they may be used to predict the coding regions from genomic DNA using exon prediction algorithms, such as FGENESH (Salamov and Solovyev, 2000, Genome Res. 10:516-522). In addition, one of ordinary skill in the art would understand how to obtain the unigene cluster identification and sequence information according to the exemplar accession numbers provided in Tables 1-16. (see, http://www.ncbi.nlm.nih.gov/UniGene/).

TABLE1: shows genes, including expression sequence tags, differentially expressed in prostate tumor tissue compared to normal tissue as analyzed using the Affymetrix/Eos Hu01 GeneChip array. Shown are the relative amounts of each gene expressed in prostate tumor samples and various normal tissue samples showing the highest expression of the gene.

Pkey:	Unique Eos probeset identifier number
ExAccn:	Exemplar Accession number, Genbank accession number
UnigeneID:	Unigene number
Unigene Title:	Unigene gene title
R1:	Ratio of tumor to normal body tissue

Pkey	UnigeneID	ExAccn	Unigene Title	R1
131919	Hs.272458	AA121266	ESTs	37.2
120328	Hs.290905	AA196979	ESTs; Weakly similar to (defline not ava	32.6
105201	Hs.31412	AA195626	ESTs	30.1
101486	Hs.1852	M24902	acid phosphatase; prostate	25.2
119073	Hs.279477	R32894	ESTs	24.8
133428	Hs.183752	M34376	microseminoprotein; beta-	23.8
128180	Hs.171995	AA595348	kallikrein 3; (prostate specific antigen	21.4
104080	Hs.57771	AA402971	Homo sapiens mRNA for serine protease (T	18.9
127537	Hs.162859	AA569531	ESTs	18.6
131665	Hs.30343	R22139	ESTs	17.4
101050	Hs.1832	K01911	neuropeptide Y	17.3
130771	Hs.1915	N48056	folate hydrolase (prostate-specific memb	17
108153	Hs.40808	AA054237	ESTs	16.9
107485	Hs.262476	W63793	S-adenosylmethionine decarboxylase 1	16.7
106155	Hs.33287	AA425309	ESTs	16.5
129534	Hs.11260	R73640	ESTs	16.4
100569	Hs.171995	HG2261-HT2351	Antigen, Prostate Specific, Alt. Splice	16
101869	Hs.181350	S39329	kallikrein 2; prostatic	15.4
135389	Hs.99872	U05237	fetal Alzheimer antigen	15
101506	Hs.62192	M27436	coagulation factor III (thromboplastin;	13.9
134374	Hs.8236	D62633	ESTs	12.7
133944	Hs.7780	AA045870	ESTs	12.5
109141	Hs.193380	AA176428	ESTs	12.3
130974	Hs.2178	X57985	H2B histone family; member Q	11.8
114768	Hs.182339	AA149007	ESTs	11.8
104394	Hs.172129	H46617	yp19h1.r1 Soares breast 3NbHBst Homo sap	11.8
125299	Hs.102720	Z39436	ESTs	11.6
104660	Hs.14846	AA007160	ESTs	11.4
100116	Hs.78045	D00654	actin; gamma 2; smooth muscle; enteric	11
131061	Hs.268744	N64328	ESTs; Moderately similar to KIAA0273 [H.	10.9
126645	126645	AI167942	Homo sapiens BAC clone RG041D11 from 7q2	10.7
135153	Hs.95420	N40141	Homo sapiens mRNA for JM27 protein; comp	10.6
107033	Hs.113314	AA599629	ESTs	10.6
118417		N66048	ESTs; Weakly similar to polymerase [H.sa	10.5
126758	Hs.293960	W37145	ESTs	10.2
115674	Hs.8364	AA406542	ESTs	10.1
134989	Hs.92381	AA236324	ESTs; Weakly similar to !!!! ALU CLASS A	10.1
107102	Hs.30652	AA609723	ESTs	10.1
116787	Hs.15641	H28581	ESTs	10.1
115719	Hs.59622	AA416997	ESTs	10
123209	Hs.203270	AA489711	ESTs	9.9
101664	Hs.121017	M60752	H2A histone family; member A	9.8
112971	Hs.83883	T17185	ESTs	9.7
102519	Hs.80296	U52969	Purkinje cell protein 4	9.7
117984	Hs.106778	N51919	ESTs	9.7
105840	Hs.22209	AA398533	ESTs	9.4
129523	Hs.274509	M30894	T-cell receptor; gamma cluster	9.4
132964	Hs.167133	AA031360	ESTs	9.2
121853	Hs.98502	AA425887	ESTs	9

115764	Hs.91011	AA421562	anterior gradient 2 (Xenopus laevis; sec	8.9
119617	Hs.55999	W47380	ESTs	8.9
100552	Hs.301946	HG2167-HT2237	Protein Kinase Ht31, Camp-Dependent	8.9
105627	Hs.23317	AA281245	ESTs	8.8
101461	Hs.76422	M22430	phospholipase A2; group IIA (platelets;	8.7
131725	Hs.31146	AA456264	ESTs; Highly similar to (defline not ava	8.5
124526	Hs.293185	N62096	yz61c5.s1 Soares_multiple_sclerosis_2NbH	8.5
118528	Hs.49397	N67869	ESTs	8.2
133845	Hs.76704	T68510	ESTs	8.2
133354	Hs.334762	AA055552	ESTs; Weakly similar to KIAA0319 [H.sapi	8.1
105912	Hs.20415	AA402000	ESTs; Weakly similar to GS3786 [H.sapien	8
119018	Hs.278695	N95796	ESTs	8
100394	Hs.66052	D84276	CD38 antigen (p45)	8
114132	Hs.24192	Z38688	ESTs	7.9
116786	Hs.301527	H25836	tumor necrosis factor (ligand) superfami	7.7
106579	Hs.23023	AA456135	ESTs	7.6
128790	Hs.105700	AA291725	secreted frizzled-related protein 4	7.5
114965	Hs.72472	AA250737	ESTs	7.4
112033	Hs.22627	R43162	ESTs	7.1
102398		U42359	Human N33 protein form 1 (N33) gene, exo	7
101201	Hs.2256	L22524	matrix metalloproteinase 7 (matrilysin;	6.9
109272	Hs.288462	AA195718	ESTs	6.9
103145	Hs.169849	X66276	myosin-binding protein C; slow-type	6.9
101803	Hs.155691	M86546	pre-B-cell leukemia transcription factor	6.8
120562	Hs.302267	AA280036	ESTs; Weakly similar to W01A6.c [C.elega	6.8
109112	Hs.257924	AA169379	ESTs	6.8
109795	Hs.326416	F10707	ESTs	6.7
107532	Hs.173684	Z19643	ESTs; Weakly similar to (defline not ava	6.7
130336	Hs.171995	X07730	kallikrein 3; (prostate specific antigen	6.6
131425	Hs.26691	AA219134	ESTs	6.6
120568	Hs.16193	AA281591	Homo sapiens mRNA; cDNA DKFZp586B211 (fr	6.6
132902	Hs.59838	AA490969	ESTs	6.6
125674	Hs.323378	W28078	H.sapiens mRNA for transmembrane protein	6.6
133724	Hs.75746	U07919	aldehyde dehydrogenase 6	6.5
130343	Hs.278628	AA490262	ESTs; Moderately similar to APXL gene pr	6.5
120215	Hs.108787	Z41050	Homo sapiens Mcd4p homolog mRNA; complet	6.5
129215	Hs.126085	AA176867	ESTs	6.5
131881	Hs.3383	AA010163	upstream regulatory element binding prot	6.5
133376	Hs.7232	T23670	ESTs	6.4
105376	Hs.8768	AA236559	ESTs; Weakly similar to neuronal thread	6.4
104674	Hs.26289	AA009527	ESTs	6.4
100727	Hs.334786	X07290	Human HF.12 gene mRNA	6.3
130150	Hs.15113	AF000573	homogentisate 1;2-dioxygenase (homogenti	6.3
121770	Hs.278428	AA421714	Homo sapiens mRNA for KIAA0896 protein;	6.3
123475	Hs.250528	AA599267	ESTs; Weakly similar to ANKYRIN; BRAIN V	6.3
133061	Hs.296638	AB000584	prostate differentiation factor	6.3
116429	Hs.279923	AA609710	ESTs; Weakly similar to similar to GTP-b	6.2
101233	Hs.878	L29008	sorbitol dehydrogenase	6.2
104691	Hs.37744	AA011176	ESTs	6.2
127248		AA325029	EST27953 Cerebellum II Homo sapiens cDNA	6.2
127775	Hs.179902	H04106	ESTs; Weakly similar to (defline not ava	6.2
105500	Hs.222399	AA256485	ESTs	6.1
131463	Hs.2714	X74142	forkhead (Drosophila)-like 1	6.1
132116	Hs.40289	AA234767	ESTs	6
130828	Hs.203213	AA053400	ESTs	5.9
115357	Hs.72988	AA281793	ESTs	5.8
105496	Hs.301997	AA256323	ESTs	5.7
116334	Hs.48948	AA491457	ESTs	5.7
107968	Hs.61539	AA034020	ESTs	5.7
120132	Hs.125019	Z38839	ESTs; Weakly similar to !!!! ALU SUBFAMI	5.6
106375	Hs.289072	AA443993	ESTs	5.6
132550	Hs.170195	AA029597	bone morphogenetic protein 7 (osteogenic	5.6
124777	Hs.140237	R41933	ESTs; Weakly similar to neuronal thread	5.6
100311	Hs.337616	D50640	phosphodiesterase 3B; cGMP-inhibited	5.6
101791	Hs.62354	M83822	Human beige-like protein (BGL) mRNA; par	5.5
117698	Hs.45107	N41002	ESTs	5.5
132387	Hs.281434	R70914	heat shock 70kD protein 1	5.5
122041	Hs.98732	AA431407	Homo sapiens Chromosome 16 BAC clone CIT	5.5
133723	Hs.262476	AA088851	S-adenosylmethionine decarboxylase 1	5.5

113938		W81598	ESTs	5.4
133015	Hs.246315	AA047036	ESTs	5.4
125745	Hs.75722	AI283493	ribophorin II	5.4
107295	Hs.80120	T34527	UDP-N-acetyl-alpha-D-galactosamine:polyp	5.4
108186	Hs.7780	AA056482	ESTs	5.3
100184	Hs.21223	D17408	calponin 1; basic; smooth muscle	5.3
104466	Hs.326392	N25110	Human guanine nucleotide exchange factor	5.3
104033	Hs.98944	AA365031	ESTs	5.3
110844	Hs.167531	N31952	ESTs; Weakly similar to (defline not ava	5.3
129056	Hs.108336	H70627	ESTs; Weakly similar to !!!! ALU SUBFAMI	5.3
102805	Hs.25351	U90304	iroquois-class homeodomain protein	5.3
133493	Hs.194369	AA284143	Homo sapiens chromosome 1 atrophin-1 rel	5.3
129184	Hs.109201	W26769	ESTs; Highly similar to (defline not ava	5.2
134158	Hs.79428	U15174	BCL2/adenovirus E1B 19kD-interacting pro	5.2
107240	Hs.159872	D59368	ESTs	5.2
104787		AA027317	ESTs; Weakly similar to !!!! ALU SUBFAMI	5.2
123527	Hs.108327	AA608679	damage-specific DNA binding protein 1 (1	5.2
116646	Hs.194228	F03048	ESTs; Moderately similar to !!!! ALU SUB	5.2
101448	Hs.195850	M21389	keratin 5 (epidermolysis bullosa simplex	5.1
116188	Hs.184598	AA464728	ESTs; Weakly similar to !!!! ALU SUBFAMI	5.1
126259	Hs.281428	Z21472	ESTs; Moderately similar to !!!! ALU SUB	5.1
105921	Hs.169119	AA402613	ESTs	5.1
103375	Hs.54416	X91868	sine oculis homeobox (Drosophila) homolo	5.1
128871	Hs.106778	AA400271	ESTs; Highly similar to (defline not ava	5.1
112681	Hs.148932	R87331	ESTs; Moderately similar to semaphorin V	5.1
105784	Hs.226434	AA350771	ESTs	5.1
116238	Hs.47144	AA479362	ESTs	5
102913	Hs.80342	X07696	keratin 15	5
103011	Hs.326035	X52541	early growth response 1	5
126023		H58881	yr36d09.r1 Soares fetal liver spleen 1NF	5
103709	Hs.13804	AA037316	ESTs	5
118981	Hs.39288	N93839	ESTs; Weakly similar to !!!! ALU SUBFAMI	5
134807	Hs.89732	X78932	zinc finger protein 273	5
100079	Hs.23311	AB002365	Human mRNA for KIAA0367 gene; partial cd	4.9
132047	Hs.3796	D83492	EphB6	4.9
132880	Hs.177537	AA444369	ESTs	4.9
124049	Hs.74519	F10523	primase; polypeptide 2A (58kD)	4.8
133330	Hs.71119	U42360	Human N33 mRNA; complete cds	4.8
104776		AA026349	ESTs	4.8
122593	Hs.128749	AA453310	Homo sapiens alpha-methylacyl-CoA racema	4.8
103912	Hs.143087	AA251078	ESTs	4.8
113961	Hs.26009	W86307	Homo sapiens mRNA for KIAA0860 protein;	4.8
105288	Hs.3585	AA233168	ESTs; Weakly similar to coded for by C.	4.8
135035	Hs.284186	H89575	ESTs	4.8
104144	Hs.183390	AA447439	ESTs; Weakly similar to ZINC FINGER PROT	4.8
129389	Hs.268126	AA621604	ESTs	4.8
125982		R98091	RAE1 (RNA export 1; S.pombe) homolog	4.8
125162	Hs.26243	W44682	ESTs	4.8
103023	Hs.117950	X53793	multifunctional polypeptide similar to S	4.7
129735		W80701	ESTs; Weakly similar to HERV-E envelope	4.7
104479	Hs.106390	N36040	ESTs	4.7
103731		AA070545	zm7c3.r1 Stratagene neuroepithelium (#93	4.7
126575	Hs.127602	W72416	ESTs	4.7
124578	Hs.231500	N68321	Human glucose transporter-like protein-l	4.7
130617	Hs.1674	M90516	glutamine-fructose-6-phosphate transamin	4.7
116752	Hs.91622	H06373	Homo sapiens clone 24456 mRNA sequence	4.7
100279	Hs.82007	D42084	Human mRNA for KIAA0094 gene; partial cd	4.7
126288	Hs.89576	AI479264	ESTs	4.7
131836	Hs.32990	AA610086	ESTs	4.7
106717	Hs.239489	AA465093	TIA1 cytotoxic granule-associated RNA-bi	4.7
114542	Hs.91011	AA055768	ESTs	4.6
103806		AA130614	zo1f2.r1 Stratagene neuroepithelium NT2R	4.6
130529		AA173238	small inducible cytokine A5 (RANTES)	4.6
115675	Hs.82065	AA406546	ESTs	4.6
111386	Hs.293798	N95326	ESTs	4.6
106503	Hs.29679	AA452411	ESTs	4.6
119943	Hs.14158	W86835	copine III	4.6
104459	Hs.100070	M91493	EST	4.6
100774	Hs.89603	HG371-HT1063	Mucin 1, Epithelial, Alt. Splice 6	4.6

100652	Hs.142653	HG2825-HT2949	Ret Transforming Gene	4.6
132015	Hs.3731	D11900	ESTs	4.6
126086		H70975	yr73g01.r1 Soares fetal liver spleen 1NF	4.6
130888	Hs.173094	F03819	ESTs	4.6
106390	Hs.20166	AA446964	Prostate stem cell antigen	4.6
126959		AA199853	ESTs; Moderately similar to !!!! ALU SUB	4.5
131584	Hs.29117	X91648	H.sapiens mRNA for pur alpha extended 3'	4.5
104838	Hs.20953	AA039481	ESTs	4.5
125661		R50319	ESTs	4.5
103171	Hs.234726	X68733	alpha-1-antichymotrypsin	4.5
103928	Hs.199160	AA280085	ESTs	4.5
102899	Hs.75730	X06272	signal recognition particle receptor ('d	4.5
100892	Hs.180789	HG4557-HT4962	Small Nuclear Ribonucleoprotein U1, 1snr	4.5
106167	Hs.7956	AA425906	ESTs	4.5
129404	Hs.317584	AA172056	ESTs	4.5
106990	Hs.24758	AA521354	ESTs	4.5
132316	Hs.44566	U28831	Human protein immuno-reactive with anti-	4.4
132056	Hs.38176	T89386	Homo sapiens mRNA for KIAA0606 protein;	4.4
133718	Hs.198760	X15306	neurofilament; heavy polypeptide (200kD)	4.4
101470	Hs.1846	M22898	tumor protein p53 (Li-Fraumeni syndrome)	4.4
131904	Hs.284296	AA143019	ESTs; Highly similar to surface 4 integr	4.4
105804	Hs.22514	AA383142	ESTs	4.4
122861	Hs.119394	AA464428	ESTs	4.4
111336	Hs.29894	N79565	ESTs	4.4
121944	Hs.98518	AA429278	ESTs	4.4
134401	Hs.211577	AA243746	ESTs; Highly similar to CG1 protein [H.s	4.4
126458	Hs.288969	AA815252	ESTs; Weakly similar to !!!! ALU SUBFAMI	4.4
133435	Hs.323966	T23983	ESTs; Moderately similar to !!!! ALU SUB	4.4
105178	Hs.21941	AA187490	ESTs	4.3
127315		AA640834	nr27b06.r1 NCI_CGAP_Pr3 Homo sapiens cDN	4.3
132645	Hs.54424	X87870	H.sapiens mRNA for hepatocyte nuclear fa	4.3
116162	Hs.282990	AA461487	ESTs; Weakly similar to F52C12.2 [C.eleg	4.3
118040	Hs.47567	N52876	EST	4.3
130008	Hs.278427	M31423	cerebellar degeneration-related protein	4.3
126607	Hs.114688	W87424	ESTs	4.3
123061	Hs.105130	AA482030	EST	4.3
109391	Hs.184245	AA219699	ESTs	4.3
109175		AA160496	ESTs	4.3
127003	Hs.173540	AA550806	ESTs; Weakly similar to (defline not ava	4.3
102547	Hs.46638	U57911	chromosome 11 open reading frame 8	4.3
134208	Hs.79993	U88871	peroxisomal biogenesis factor 7	4.3
104258	Hs.5462	AF007216	solute carrier family 4; sodium bicarbon	4.3
130759	Hs.18946	AA094720	ESTs; Weakly similar to (defline not ava	4.3
132160	Hs.295923	AA281770	seven in absentia (Drosophila) homolog 1	4.3
135062	Hs.93872	AA174183	ESTs	4.3
126510	Hs.334762	R49702	ESTs; Weakly similar to KIAA0319 [H.sapi	4.2
122055	Hs.98747	AA431732	EST	4.2
133136	Hs.6574	AF007165	suppressin (nuclear deformed epidermal a	4.2
109890	Hs.20843	H04649	ESTs	4.2
133294	Hs.69997	R79723	H.sapiens mRNA for translin associated z	4.2
134436	Hs.83190	S80437	fatty acid synthase (3' region) [human,	4.2
107375	Hs.251064	U88573	NBR2	4.2
122223	Hs.27413	AA436158	ESTs	4.2
103044	Hs.248210	X55777	H.sapiens Mahlavu hepatocellular carcino	4.2
120125	Hs.59815	W99362	EST	4.2
128969	Hs.283978	T65327	ESTs; Highly similar to (defline not ava	4.2
129637	Hs.1179	D90359	TATA box binding protein (TBP)-associate	4.2
106566		AA455921	ESTs; Weakly similar to !!!! ALU SUBFAMI	4.2
112605	Hs.29852	R79220	ESTs	4.2
103364	Hs.279929	X90872	H.sapiens mRNA for gp25L2 protein	4.2
132811	Hs.57419	U25435	transcriptional repressor	4.2
126570	Hs.326292	T79274	ESTs	4.2
116298	Hs.94109	AA489046	ESTs	4.2
103024	Hs.105938	X53961	lactotransferrin	4.1
129133	Hs.109850	R56728	yg95c6.r1 Soares infant brain 1NIB Homo	4.1
133167	Hs.6641	N98707	kinesin family member 5C	4.1
126871	Hs.14051	AA351779	ESTs	4.1
132333	Hs.45032	AA192157	ESTs	4.1
107376	Hs.327179	U90545	solute carrier family 17 (sodium phospha	4.1

128517	Hs.100861	AA280617	ESTs; Weakly similar to p60 katanin [H.s	4.1
130555	Hs.116774	AA450324	ESTs	4.1
105765	Hs.24183	AA343514	ESTs	4.1
126529	Hs.26369	AA133237	ESTs	4.1
125928	Hs.181889	H29730	ESTs	4.1
117280	Hs.172129	N22107	ESTs; Moderately similar to !!!! ALU SUB	4.1
100234	Hs.3085	D29677	KIAA0054 gene product	4.1
100959	Hs.118127	J00073	actin; alpha; cardiac muscle	4.1
107130	Hs.12913	AA620582	ESTs; Weakly similar to (defline not ava	4.1
105035	Hs.8659	AA128486	ESTs	4.1
126735	Hs.226795	AA808949	glutathione S-transferase pi	4.1
113056	Hs.8036	T26471	ESTs; Moderately similar to !!!! ALU SUB	4
102460	Hs.211582	U48959	Homo sapiens myosin light chain kinase (	4
106968	Hs.26813	AA504631	ESTs; Weakly similar to (defline not ava	4
123107	Hs.104207	AA486071	ESTs	4
127256	Hs.267967	AA327550	ESTs; Weakly similar to !!!! ALU SUBFAMI	4
105329	Hs.22862	AA234561	ESTs	4
115504	Hs.42736	AA291946	ESTs	4
120726	Hs.97293	AA293656	ESTs	4
103576	Hs.94560	Z26317	desmoglein 2	4
127889	Hs.144941	AI147408	ESTs	4
106394	Hs.25320	AA447223	ESTs	4
128046		AA873285	ESTs	4
103391	Hs.114366	X94453	pyrroline-5-carboxylate synthetase (glut	4
106448	Hs.27004	AA449455	ESTs	4
126513	Hs.86276	W27601	ESTs; Moderately similar to (defline not	4
129593	Hs.98314	AA487015	ESTs; Weakly similar to !!!! ALU SUBFAMI	3.9
110151	Hs.31608	H18836	ESTs	3.9
105344	Hs.8645	AA235303	ESTs	3.9
104791	Hs.301871	AA029046	ESTs	3.9
123442	Hs.111496	AA598803	ESTs	3.9
127800	Hs.79428	AA521047	BCL2/adenovirus E1B 19kD-interacting pro	3.9
114555	Hs.167904	AA058594	ESTs	3.9
122138	Hs.163960	AA435549	ESTs	3.9
129565	Hs.198726	X77777	vasoactive intestinal peptide receptor 1	3.9
103471	Hs.75216	Y00815	protein tyrosine phosphatase; receptor t	3.9
133908	Hs.325474	M83216	caldesmon 1	3.9
105635	Hs.301985	AA281508	ESTs	3.9
134285	Hs.81086	AA460012	solute carrier family 22 (organic cation	3.9
134125	Hs.50421	R38102	KIAA0203 gene product	3.9
125628	Hs.241493	AA418069	natural killer-tumor recognition sequenc	3.9
103695	Hs.186600	AA018758	ESTs	3.9
100642	Hs.182183	HG2743-HT3926	Caldesmon 1, Alt. Splice 6, Non-Muscle	3.9
104334	Hs.78771	D82614	ESTs	3.9
110242	Hs.19978	H26417	ESTs	3.9
125298	Hs.289008	Z39255	ESTs	3.9
104060	Hs.303193	AA397968	zt87a9.r1 Soares_testis_NHT Homo sapiens	3.9
105823	Hs.293960	AA398197	ESTs	3.9
126499	Hs.110445	AA315671	ESTs; Moderately similar to unknown [M.m	3.9
130752	Hs.18895	D50927	KIAA0137 gene product	3.8
123494	Hs.112110	AA599786	ESTs	3.8
104846	Hs.32478	AA040154	ESTs	3.8
106921	Hs.71721	AA142913	ESTs	3.8
115506	Hs.45207	AA292537	ESTs	3.8
100452	Hs.241552	D87742	Human mRNA for KIAA0268 gene; partial cd	3.8
104454	Hs.129228	M84443	galactokinase 2	3.8
108730	Hs.102859	AA126254	ESTs	3.8
131223	Hs.24427	AA247788	ESTs; Highly similar to (defline not ava	3.8
104784	Hs.269228	AA027055	ESTs	3.8
104946	Hs.73848	AA069549	ESTs	3.8
106932	Hs.9394	AA495926	ESTs	3.8
101724	Hs.620	M69225	bullous pemphigoid antigen 1 (230/240kD)	3.8
106140	Hs.14912	AA424524	Homo sapiens mRNA for KIAA0286 gene; par	3.8
128135	Hs.269721	AA913491	ESTs	3.8
120030	Hs.58694	W92051	ESTs	3.8
126457	Hs.50382	AA007489	zh98g04.r1 Soares_fetal_liver_spleen_1NF	3.8
123917	Hs.112969	AA621311	EST	3.7
110714	Hs.17752	H95978	Homo sapiens phosphatidylserine-specific	3.7
130577	Hs.162	M35410	insulin-like growth factor binding prote	3.7

117667	Hs.44708	N39214	ser-Thr protein kinase related to the my	3.7
126104	Hs.39712	N77278	ESTs; Weakly similar to BONE/CARTILAGE P	3.7
100379	Hs.278721	D82060	Homo sapiens mRNA for membrane protein w	3.7
115646	Hs.305971	AA404352	ESTs	3.7
125792	Hs.193700	AI005388	ESTs; Moderately similar to !!!! ALU SUB	3.7
102162	Hs.1592	U18291	CDC16 (cell division cycle 16; S. cerevi	3.7
126530	Hs.183475	AA504343	ESTs; Moderately similar to !!!! ALU SUB	3.7
119940	Hs.272531	W86779	EST	3.7
110769	Hs.23837	N22222	yw34b06.s1 Morton Fetal Cochlea Homo sap	3.7
132914	Hs.60293	AA496037	ESTs	3.7
113594	Hs.15683	T92030	ESTs	3.7
103702	Hs.279952	AA027793	ESTs; Highly similar to (defline not ava	3.7
130780	Hs.19347	AA248406	ESTs	3.7
123288	Hs.291025	AA495836	EST	3.7
120691	Hs.22380	AA291173	ESTs	3.7
103153	Hs.75295	X66534	guanylate cyclase 1; soluble; alpha 3	3.7
129201	Hs.109390	H19969	ESTs	3.7
114798	Hs.54900	AA159181	ESTs	3.7
126801	Hs.7337	AA512902	ESTs	3.7
105503	Hs.31707	AA256616	ESTs	3.7
104260	Hs.194283	AF008192	Homo sapiens putative GR6 protein (GR6)	3.7
125980	Hs.35699	R97219	ESTs	3.7
123255	Hs.105273	AA490890	ESTs	3.6
103862	Hs.6363	AA206625	ESTs	3.6
100696	Hs.121686	HG3162-HT3339	Transcription Factor IIa	3.6
134917	Hs.166994	X87241	FAT tumor suppressor (Drosophila) homolo	3.6
103520		Y10511	H.sapiens mRNA for CD176 protein	3.6
113778	Hs.302738	W15263	ESTs	3.6
101838	Hs.75511	M92934	connective tissue growth factor	3.6
113702		T97307	ESTs; Moderately similar to !!!! ALU SUB	3.6
118201	Hs.48428	N59800	EST	3.6
116519	Hs.68554	C20760	EST	3.6
105866	Hs.22983	AA400517	ESTs; Moderately similar to UDP-GLUCOSE:	3.6
106709	Hs.170291	AA464696	ESTs	3.6
127858	Hs.27973	AA806365	oc26h07.s1 NCI_CGAP_GCB1 Homo sapiens cD	3.6
101964		S81578	dioxin-responsive gene {putative polyade	3.6
105508	Hs.326416	AA256680	ESTs	3.6
116844	Hs.337434	H64938	ESTs	3.6
105372	Hs.142296	AA236481	ESTs	3.6
100745	Hs.144630	HG3510-HT3704	V-Erba Related Ear-3 Protein	3.6
127521	Hs.164018	AA809982	ESTs	3.6
110758	Hs.274265	N21365	talin	3.6
107307	Hs.44155	T52099	creatine kinase; mitochondrial 2 (sarcom	3.6
133200	Hs.183639	AA432248	ESTs	3.6
114774	Hs.184325	AA150043	ESTs	3.6
120265	Hs.270696	AA173759	ESTs; Moderately similar to !!!! ALU SUB	3.6
134359	Hs.199067	M34309	v-erb-b2 avian erythroblastic leukemia v	3.6
116250	Hs.44829	AA480975	ESTs; Moderately similar to !!!! ALU SUB	3.6
106313	Hs.35841	AA436459	nuclear factor I/X (CCAAT-binding transc	3.6
131898	Hs.279780	N52232	ESTs	3.6
133444	Hs.73793	M27281	vascular endothelial growth factor	3.6
128232	Hs.334641	H06296	ESTs	3.6
135357	Hs.79572	AA235803	ESTs	3.5
457951		AI369384	arylsulfatase D	3.5
108407		AA075519	zm87h9.s1 Stratagene ovarian cancer (#93	3.5
126659		T16245	a disintegrin and metalloproteinase doma	3.5
104189	Hs.301804	AA485805	ESTs	3.5
125956	Hs.129014	N53276	ESTs	3.5
103026	Hs.79386	X54162	Human mRNA for a 64 Kd autoantigen expre	3.5
133011	Hs.171921	AA042990	sema domain; immunoglobulin domain (Ig);	3.5
131379	Hs.26176	R49035	ESTs	3.5
126742	Hs.169359	H64106	yr57e06.r1 Soares fetal liver spleen 1NF	3.5
105560	Hs.306915	AA262783	ESTs	3.5
118472	Hs.42179	N66818	ESTs	3.5
105623	Hs.30127	AA280895	ESTs; Highly similar to !!!! ALU SUBFAMI	3.5
120262	Hs.145807	AA172076	ESTs; Moderately similar to !!!! ALU SUB	3.5
105027	Hs.26771	AA126472	ESTs	3.5
130760	Hs.18953	AA128997	phosphodiesterase 9A	3.5
117473	Hs.155560	N30157	ESTs	3.5

102663	Hs.168075	U70322	karyopherin (importin) beta 2	3.5
126349	Hs.13531	AA442868	ESTs; Weakly similar to (defline not ava	3.5
132154	Hs.41119	N67179	ESTs	3.5
131689	Hs.30696	AA599653	transcription factor-like 5 (basic helix	3.5
127862	Hs.163191	AA765305	EST	3.5
126995	Hs.189810	W26950	Human DNA sequence from PAC 388M5 on chr	3.5
119071		R31180	ESTs	3.5
103941	Hs.96593	AA282978	ESTs	3.5
110721	Hs.31319	H97678	ESTs	3.5
126586	Hs.43086	AA011247	ESTs	3.5
103106	Hs.1857	X62025	phosphodiesterase 6G; cGMP-specific; rod	3.5
116357	Hs.90797	AA504806	Homo sapiens clone 23620 mRNA sequence	3.5
105309	Hs.4104	AA233790	ESTs	3.5
130796	Hs.19525	R39390	ESTs	3.5
109101	Hs.52184	AA167708	ESTs	3.5
103134	Hs.2839	X65724	Norrie disease (pseudoglioma)	3.5
131798	Hs.301449	X86098	adenovirus 5 E1A binding protein	3.5
118535	Hs.49418	N67968	ESTs	3.5
102592	Hs.11223	U62389	Human putative cytosolic NADP-dependent	3.4
125905	Hs.6456	T69868	chaperonin containing TCP1; subunit 2 (b	3.4
109160	Hs.301997	AA179387	ESTs	3.4
105327	Hs.211593	AA234440	ESTs	3.4
106586	Hs.57787	AA456598	ESTs	3.4
122635		AA454085	EST	3.4
132413	Hs.260116	AA132969	metalloprotease 1 (pitrilysin family)	3.4
131938	Hs.34956	AA283620	ESTs	3.4
133871	Hs.182793	AA454597	ESTs	3.4
107175	Hs.292503	AA621751	ESTs; Weakly similar to KIAA0601 protein	3.4
101188	Hs.184298	L20320	cyclin-dependent kinase 7 (homolog of Xe	3.4
126422	Hs.237658	H48518	ESTs; Highly similar to apolipoprotein A	3.4
118475		N66845	ESTs; Weakly similar to !!!! ALU CLASS B	3.4
104558	Hs.88959	R56678	ESTs; Weakly similar to !!!! ALU SUBFAMI	3.4
128307	Hs.132005	AI453794	ESTs	3.4
112254	Hs.25829	R51831	ESTs	3.4
125408	Hs.89578	N72353	yv37e12.r1 Soares fetal liver spleen 1NF	3.4
109834	Hs.175955	H00804	ESTs	3.4
130844	Hs.20191	D12122	seven in absentia (Drosophila) homolog 2	3.4
127143	Hs.20843	AA533553	nj68h04.s1 NCI_CGAP_Pr10 Homo sapiens cD	3.4
135309	Hs.42500	D25984	ESTs	3.4
125724	Hs.295978	AA083407	stimulated trans-acting factor (50 kDa)	3.4
127692	Hs.187983	AI021912	ESTs	3.4
116674	Hs.92127	F04816	ESTs	3.4
134700	Hs.8868	AA481414	golgi SNAP receptor complex member 1	3.4
114846	Hs.166196	AA234929	ESTs	3.4
103649	Hs.155983	Z70219	H.sapiens mRNA for 5'UTR for unknown pro	3.4
134835	Hs.89925	L04569	calcium channel; voltage-dependent; L ty	3.4
130568	Hs.16085	AA232535	ESTs; Highly similar to (defline not ava	3.4
111331	Hs.15978	N78773	ESTs	3.4
106036	Hs.10653	AA412505	ESTs	3.4
130987	Hs.21893	R45698	ESTs	3.4
112814	Hs.35828	R98192	ESTs	3.4
127815	Hs.255015	AA876009	ob93c10.s1 NCI_CGAP_GCB1 Homo sapiens cD	3.4
100144	Hs.75616	D13643	KIAA0018 gene product	3.4
101129	Hs.247992	L10405	Homo sapiens DNA binding protein for sur	3.4
130874	Hs.20621	T08287	ESTs	3.4
106882	Hs.26994	AA489009	ESTs	3.4
103855	Hs.302267	AA195179	ESTs	3.4
125957		H45213	yo03b08.r1 Soares adult brain N2b5HB55Y	3.3
114048	Hs.146085	W94613	ESTs	3.3
109826	Hs.75354	F13702	ESTs	3.3
125355	Hs.170098	R45630	ESTs; Highly similar to KIAA0372 [H.sapi	3.3
104182	Hs.143792	AA479990	ESTs; Weakly similar to glioma amplified	3.3
100294	Hs.75454	D49396	Human mRNA for Apo1_Human (MER5(Aop1-Mou	3.3
131688	Hs.30692	U24153	p21 (CDKN1A)-activated kinase 2	3.3
116256	Hs.88201	AA481256	ESTs; Weakly similar to (defline not ava	3.3
102034	Hs.230	U05291	fibromodulin	3.3
130072	Hs.14658	R99606	Human chromosome 5q13.1 clone 5G8 mRNA	3.3
114615	Hs.159456	AA083812	ESTs; Highly similar to (defline not ava	3.3
128707	Hs.104105	AA136474	Meis (mouse) homolog 2	3.3

115048	Hs.190057	AA252668	ESTs	3.3
125862	Hs.31110	H12084	ESTs	3.3
135142	Hs.24192	R31679	ESTs	3.3
103119	Hs.2877	X63629	cadherin 3; P-cadherin (placental)	3.3
104460	Hs.62604	M91504	ESTs	3.3
100365	Hs.79284	D78611	mesoderm specific transcript (mouse) hom	3.3
131524	Hs.301804	N39152	ESTs	3.3
102165	Hs.159627	U18321	Death associated protein 3	3.3
126966	Hs.182575	R38438	solute carrier family 15 (H+/peptide tra	3.3
124839	Hs.140942	R55784	ESTs	3.3
100709	Hs.100469	HG3264-HT3441	Af-6 (Gb:U02478)	3.3
132967	Hs.61635	AA032221	Homo sapiens BAC clone RG041D11 from 7q2	3.3
102927	Hs.65114	X12876	keratin 18	3.3
132616	Hs.283558	AA386264	ESTs	3.3
125132	Hs.129781	W15495	ESTs	3.3
111225	Hs.31652	N68989	ESTs	3.3
114956	Hs.87113	AA243681	ESTs	3.3
122235	Hs.112227	AA436475	ESTs	3.3
112325	Hs.12315	R56055	ESTs	3.3
123360	Hs.178604	AA504784	ESTs	3.3
105150	Hs.155995	AA169640	Homo sapiens mRNA for KIAA0643 protein;	3.3
107391	Hs.284294	W02877	ESTs	3.3
113058	Hs.7569	T26893	EST	3.3
134371	Hs.82318	S69790	Brush-1	3.3
125669	Hs.333256	R51308	ESTs; Moderately similar to !!!! ALU SUB	3.3
111506	Hs.294105	R07726	ESTs	3.3
122974	Hs.194215	AA478625	ESTs	3.3
102369	Hs.299867	U39840	hepatocyte nuclear factor 3; alpha	3.3
120408	Hs.190151	AA235045	ESTs	3.3
117993	Hs.47402	N52039	ESTs; Weakly similar to !!!! ALU SUBFAMI	3.3
129586	Hs.11500	AA437118	ESTs	3.3
128138	Hs.126494	AI200825	ESTs	3.3
127265		AA332751	EST37214 Embryo, 8 week I Homo sapiens c	3.3
107674	Hs.41143	AA011027	Homo sapiens mRNA for KIAA0581 protein;	3.2
104866	Hs.293691	AA045342	ESTs	3.2
103427	Hs.250655	X97303	H.sapiens mRNA for Ptg-12 protein	3.2
132990	Hs.334334	AA458761	ESTs	3.2
127017	Hs.251946	AA740146	ESTs	3.2
132313	Hs.44481	U13220	forkhead (Drosophila)-like 6	3.2
106880	Hs.32425	AA488889	ESTs	3.2
107039	Hs.169780	AA599751	homologous to yeast nitrogen permease (c	3.2
120870	Hs.292581	AA357172	ESTs	3.2
107920	Hs.284207	AA027951	ESTs	3.2
104165	Hs.105116	AA459160	EST	3.2
107012	Hs.63908	AA598745	ESTs	3.2
103605	Hs.194657	Z35402	H.sapiens gene encoding E-cadherin, exon	3.2
124006	Hs.270016	D60302	ESTs	3.2
101300	Hs.74137	L40391	Homo sapiens (clone s153) mRNA fragment	3.2
101183	Hs.795	L19779	H2A histone family; member O	3.2
125596		R25698	yg44h11.r2 Soares infant brain 1NIB Homo	3.2
127261		AA661567	nu86b02.s1 NCI_CGAP_Alv1 Homo sapiens cD	3.2
120090	Hs.59554	W94591	ESTs	3.2
129393	Hs.166982	D13435	phosphatidylinositol glycan; class F	3.2
120923	Hs.97129	AA382283	ESTs	3.2
118907	Hs.274256	N91003	ESTs	3.2
111552	Hs.191185	R09411	ESTs	3.2
104431	Hs.99913	J03019	adrenergic; beta-1-; receptor	3.2
133551	Hs.278634	D63480	Human mRNA for KIAA0146 gene; partial cd	3.2
131615	Hs.192803	D14533	xeroderma pigmentosum; complementation g	3.2
126547	Hs.84072	U47732	transmembrane 4 superfamily member 3	3.2
103172	Hs.116774	X68742	integrin; alpha 1	3.2
113867	Hs.24095	W68845	ESTs	3.2
133323	Hs.70937	Z83735	H3 histone family; member K	3.2
111597	Hs.189716	R11499	ESTs	3.2
121515	Hs.104696	AA412133	ESTs	3.2
107445	Hs.6639	W28406	ESTs	3.2
106887	Hs.334335	AA489091	ESTs	3.2
123052	Hs.185766	AA481806	ESTs	3.2
107072	Hs.130760	AA609113	Homo sapiens mRNA; cDNA DKFZp586N0318 (f	3.2

	102214	Hs.32964	U23752	SRY (sex-determining region Y)-box 11	3.2
	123147		AA487961	ab11h6.s1 Stratagene lung (#93721) Homo	3.2
	125435	Hs.272138	R00940	ye87g03.r1 Soares fetal liver spleen 1NF	3.2
	116246	Hs.250646	AA479961	ESTs; Highly similar to ubiquitin-conjug	3.2
5	105169	Hs.180789	AA180321	Homo sapiens (clone S164) mRNA; 3' end o	3.2
	134001	Hs.78344	AF001548	myosin; heavy polypeptide 11; smooth mus	3.2
	124866	Hs.304389	R68571	ESTs	3.2
	133205	Hs.67619	AA089559	Homo sapiens mRNA; chromosome 1 specific	3.2
	102986	Hs.182378	X17648	colony stimulating factor 1 (macrophage)	3.2
10	101232	Hs.242894	L28997	ADP-ribosylation factor-like 1	3.1
	132906	Hs.234896	AA142857	ESTs; Highly similar to geminin [H.sapie	3.1
	104281	Hs.5669	C14290	ESTs	3.1
	123926	Hs.227933	AA621348	ESTs; Highly similar to (defline not ava	3.1
	134464	Hs.239720	N79354	ESTs; Weakly similar to Rga [D.melanogas	3.1
15	105322	Hs.16346	AA234100	ESTs	3.1
	100631	Hs.48332	HG2709-HT2805	Serine/Threonine Kinase (Gb:Z25431)	3.1
	130791	Hs.199263	AA259102	ESTs; Highly similar to (defline not ava	3.1
	131220	Hs.300855	R77200	ESTs	3.1
	113237	Hs.123642	T62857	ESTs	3.1
20	125562	Hs.98968	AI494372	ESTs	3.1
	134110	Hs.79136	U41060	Human breast cancer; estrogen regulated	3.1
	132393	Hs.47334	W85888	ESTs; Moderately similar to !!!! ALU SUB	3.1
	107439	Hs.296842	W27995	ESTs; Moderately similar to non-muscle m	3.1
	125863	Hs.40719	AA299096	Homo sapiens mRNA; cDNA DKFZp564M0916 (f	3.1
25	105811	Hs.286192	AA394121	ESTs	3.1
	129284	Hs.296141	AA104023	ESTs	3.1
	125321	Hs.178294	T86652	ESTs	3.1
	107332	Hs.183297	T87750	ESTs	3.1
	123570	Hs.109653	AA608955	ESTs	3.1
30	100384	Hs.90800	D83646	matrix metalloproteinase 16 (membrane-in	3.1
	109063	Hs.38972	AA161043	tetraspan 1	3.1
	133284	Hs.182628	U09367	zinc finger protein 136 (clone pHZ-20)	3.1
	131839	Hs.33010	H80622	Homo sapiens mRNA for KIAA0633 protein;	3.1
	117606	Hs.44698	N35115	ESTs	3.1
35	418998	Hs.287849	F13215	ESTs	3.1
	125180	Hs.103120	W58344	ESTs	3.1
	100789		HG3893-HT4163	Phosphoglucomutase 1, Alt. Splice	3.1
	126017	Hs.159440	H60487	ESTs	3.1
	132452	Hs.247324	AA005262	Homo sapiens DNA sequence from PAC 262D1	3.1
40	129077	Hs.108479	H78886	ESTs	3.1
	126563	Hs.181368	W26247	U5 snRNP-specific protein (220 kD); orth	3.1
	129650	Hs.116258	N52554	ESTs	3.1
	123465		AA599033	ESTs	3.1
	126486	Hs.152316	AA345339	EST51345 Gall bladder II Homo sapiens cD	3.1
45	126460	Hs.167031	W01616	za36d05.r1 Soares fetal liver spleen 1NF	3.1
	118697	Hs.43234	N72094	ESTs	3.1
	103860	Hs.38057	AA203742	ESTs	3.1
	127968	Hs.124347	AA971439	ESTs	3.1
	124984	Hs.223241	T47566	yb15c11.s1 Stratagene placenta (#937225)	3.1
50	103903	Hs.15220	AA249334	j312.seq.F Human fetal heart, Lambda ZAP	3.1
	106697	Hs.22242	AA463737	ESTs	3.1
	130892	Hs.20993	AA442604	ESTs; Weakly similar to Ydr374cp [S.cere	3
	114032	Hs.35014	W92779	ESTs	3
	128835	Hs.106390	W15528	ESTs	3
55	103667	Hs.247815	Z80788	H.sapiens H4/I gene	3
	126264	Hs.250614	N42897	yy13h06.r1 Soares melanocyte 2NbHM Homo	3
	132626	Hs.21275	D25755	ESTs	3
	131107	Hs.75354	N87590	ESTs	3
	126780	Hs.5811	R12421	ESTs	3
60	127363	Hs.22116	AA307744	Homo sapiens Cdc14B1 phosphatase mRNA; c	3
	103690	Hs.84063	AA016186	ESTs	3
	102589	Hs.8867	U62015	Homo sapiens Cyr61 mRNA, complete cds	3
	125144	Hs.24336	W37999	ESTs	3
	132977	Hs.301404	U28686	RNA binding motif protein 3	3
65	120714	Hs.146170	AA292689	ESTs	3
	101038	Hs.79411	J05249	replication protein A2 (32kD)	3
	102856	Hs.248177	X00090	Human histone H3 gene	3
	105516	Hs.30738	AA257971	ESTs	3
	131137	Hs.33287	U85193	nuclear factor I/B	3

127221	Hs.241551	AI354332	ESTs	3
411888	Hs.24104	R26708	ESTs	3
131684	Hs.3066	U26174	granzyme K (serine protease; granzyme 3;	3
100629	Hs.21291	HG2706-HT2802	Serine/Threonine Kinase (Gb:Z25428)	3
119944	Hs.58915	W86838	EST	3
113801	Hs.118281	W38418	zinc finger protein 266	3
133780	Hs.76152	M14219	decorin	3
104690	Hs.14449	AA010889	ESTs	3
126371	Hs.304139	N57645	EST	3
127635	Hs.116346	AA766903	ESTs	3
128434	Hs.143880	AI190914	ESTs	3
435761	Hs.187555	AA701941	ESTs	3
125025	Hs.50748	T71561	ESTs	3
124940	Hs.103804	R99599	heterogeneous nuclear ribonucleoprotein	3
128742	Hs.251531	D00763	proteasome (prosome; macropain) subunit;	3
107147	Hs.10450	AA621125	Homo sapiens chromosome 2; 10 repeat reg	3
112068	Hs.22545	R43910	ESTs	3
105346	Hs.263727	AA235465	ESTs; Moderately similar to !!!! ALU SUB	3
130972	Hs.21739	AA370302	Homo sapiens mRNA; cDNA DKFZp586I1518 (f	3
131230	Hs.274407	AA149987	thymus specific serine peptidase	3
133743	Hs.75847	N79435	ESTs	3
127402	Hs.227949	AA358869	ESTs; Highly similar to SEC13-RELATED PR	3
117483	Hs.44189	N30426	ESTs	3
123659	Hs.112699	AA609368	ESTs	3
103963	Hs.63290	AA298588	EST114219 HSC172 cells II Homo sapiens c	3
103795	Hs.7367	AA112222	ESTs; Moderately similar to (defline not	3
115092	Hs.80975	AA255903	CD39-like 4	2.9
134831	Hs.89890	S72370	pyruvate carboxylase	2.9
128579	Hs.101810	AA093378	ESTs; Weakly similar to !!!! ALU SUBFAMI	2.9
134193	Hs.7980	F09570	ESTs	2.9
123522	Hs.112575	AA608577	ESTs	2.9
107109	Hs.32793	AA609943	ESTs	2.9
134694	Hs.88556	D50405	histone deacetylase 1	2.9
134399	Hs.82689	H99801	tumor rejection antigen (gp96) 1	2.9
134632	Hs.174139	AA398710	H. sapiens RNA for CLCN3	2.9
106683	Hs.14512	AA461495	ESTs	2.9
108555		AA084963	zn13e12.s1 Stratagene hNT neuron (#93723	2.9
100953	Hs.2110	HG945-HT945	Nucleic Acid-Binding Protein (Gb:L12693)	2.9
130597	Hs.16492	AA173998	ESTs; Weakly similar to weakly similar t	2.9
101813	Hs.139226	M87338	replication factor C (activator 1) 2 (40	2.9
106636	Hs.286	AA459950	ESTs	2.9
129109	Hs.108708	AA491295	calcium/calmodulin-dependent protein kin	2.9
125819	Hs.251871	AA044840	stromal cell-derived factor 1	2.9
106282	Hs.9857	AA433946	ESTs; Weakly similar to (defline not ava	2.9
100386	Hs.301636	D83703	peroxisomal biogenesis factor 6	2.9
114546	Hs.98074	AA056263	ESTs; Moderately similar to !!!! ALU SUB	2.9
105914	Hs.9701	AA402224	Homo sapiens growth arrest and DNA-damag	2.9
108552		AA084912	zn11c7.s1 Stratagene hNT neuron (#937233	2.9
126505	Hs.190057	W26894	16a11 Human retina cDNA randomly primed	2.9
134098	Hs.79086	X06323	Human MRL3 mRNA for ribosomal protein L3	2.9
129721	Hs.211539	L19161	eukaryotic translation initiation factor	2.9
100076	Hs.277422	AB000897	Homo sapiens mRNA for cadherin FIB3, par	2.9
117466	Hs.44104	N29862	ESTs	2.9
106335	Hs.36688	AA437258	ESTs; Moderately similar to WAP four-dis	2.9
134510	Hs.250870	U25265	protein kinase; mitogen-activated; kinas	2.9
105835	Hs.32995	AA398412	ESTs	2.9
106611	Hs.26267	AA458904	ESTs; Weakly similar to torsinA [H.sapie	2.9
134087	Hs.173824	U51166	thymine-DNA glycosylase	2.9
100641	Hs.182183	HG2743-HT2846	Caldesmon 1, Alt. Splice 4, Non-Muscle	2.9
104602		R86920	ESTs	2.9
117203	Hs.42738	H99799	ESTs	2.9
131889	Hs.34073	AA401912	BH-protocadherin (brain-heart)	2.9
101707	Hs.155212	M65131	methylmalonyl Coenzyme A mutase	2.9
115271	Hs.5724	AA279422	ESTs	2.9
125812	Hs.287912	H73420	lectin; mannose-binding; 1	2.9
110740	Hs.19762	H99675	ESTs	2.9
103406	Hs.285728	X95677	H.sapiens mRNA for ArgBPIB protein	2.9
104577	Hs.132390	R71539	ESTs	2.9
102772	Hs.161002	U83115	absent in melanoma 1	2.9

	131710	Hs.30985	AA233225	ESTs; Highly similar to (defline not ava	2.9
	125231	Hs.268903	W84714	ESTs	2.9
	127380	Hs.15535	AI417137	Homo sapiens clone 24582 mRNA sequence	2.9
	104229	Hs.61289	AB002346	inositol phosphate 5'-phosphatase 2 (syn	2.9
5	126600	Hs.191385	AA699949	ESTs	2.9
	125175	Hs.303030	W52355	EST	2.9
	103849	Hs.34578	AA187045	ESTs; Weakly similar to !!!! ALU SUBFAMI	2.9
	102126	Hs.78961	U14575	protein phosphatase 1; regulatory (inhib	2.9
	124906	Hs.107815	R87647	ESTs	2.9
10	131148	Hs.303125	C00038	ESTs	2.9
	123158	Hs.218329	AA488658	heat shock 70kD protein 1	2.9
	133667	Hs.75462	U72649	Human BTG2 (BTG2) mRNA; complete cds	2.9
	105182	Hs.18271	AA191014	ESTs; Weakly similar to Ydr372cp [S.cere	2.9
	133968	Hs.232068	D15050	Human mRNA for transcription factor AREB	2.9
15	117425	Hs.336901	N27154	ESTs	2.9
	111087	Hs.37637	N59645	ESTs	2.9
	129641	Hs.11805	N66066	ESTs	2.9
	128639	Hs.102897	N91246	ESTs	2.9
	133209	Hs.79265	AA114183	ESTs; Moderately similar to glutamate py	2.9
20	135154	Hs.267812	AA126433	sorting nexin 4	2.9
	126838	Hs.279609	AA858097	pigment epithelium-derived factor	2.9
	103803	Hs.106149	AA127696	ESTs	2.9
	102139	Hs.2128	U15932	dual specificity phosphatase 5	2.9
	128104		AA971000	op67g11.s1 Soares_NFL_T_GBC_S1 Homo sapi	2.8
25	127834	Hs.337631	AA761415	nz22d08.s1 NCI_CGAP_GCB1 Homo sapiens cD	2.8
	133101	Hs.180952	AA488230	ESTs	2.8
	127250	Hs.217916	AI023717	ESTs	2.8
	135063	Hs.93883	D10537	myelin protein zero (Charcot-Marie-Tooth	2.8
	126323	Hs.68644	N45014	yy80g06.r1 Soares_multiple_sclerosis_2Nb	2.8
30	121873	Hs.145696	AA426270	ESTs	2.8
	122090	Hs.98684	AA432141	ESTs	2.8
	118728	Hs.322645	N73705	ESTs	2.8
	135400	Hs.99915	M23263	androgen receptor (dihydrotestosterone r	2.8
	125278	Hs.129998	W93523	ESTs	2.8
35	124387	Hs.109019	N27637	ESTs	2.8
	124803	Hs.12186	R45480	cyclin K	2.8
	H45968	Hs.32149	H45968	ESTs	2.8
	104261	Hs.5409	AF008442	RNA polymerase I subunit	2.8
	105366	Hs.282093	AA236356	ESTs	2.8
40	106070	Hs.5957	AA417761	Homo sapiens clone 24416 mRNA sequence	2.8
	131356	Hs.25960	M13241	v-myc avian myelocytomatosis viral relat	2.8
	112009	Hs.26255	R42714	EST	2.8
	133199	Hs.250175	AA609773	Homo sapiens clone 23904 mRNA sequence	2.8
	110379	Hs.33130	H44825	ESTs	2.8
45	103890	Hs.72085	AA236843	ESTs; Weakly similar to unknown [S.cerev	2.8
	128152		R20353	yg20f10.r1 Soares infant brain 1NIB Homo	2.8
	107008	Hs.23740	AA598710	ESTs	2.8
	135243	Hs.97101	AA215333	ESTs	2.8
	103058	Hs.184510	X57348	stratifin	2.8
50	132020	Hs.293845	AA428990	ESTs	2.8
	116354	Hs.292566	AA504262	ESTs	2.8
	125867	Hs.12372	H98141	ESTs	2.8
	120603	Hs.98541	AA282787	ESTs; Highly similar to (defline not ava	2.8
	115119	Hs.46847	AA256524	Human DNA sequence from clone 30M3 on ch	2.8
55	133865	Hs.170290	F09315	discs; large (Drosophila) homolog 5	2.8
	109415	Hs.110826	AA227219	Homo sapiens CAGF9 mRNA; partial cds	2.8
	128687	Hs.23767	Z38910	ESTs	2.8
	109984	Hs.10299	H09594	ESTs; Moderately similar to !!!! ALU SUB	2.8
	133179	Hs.66731	U81599	homeo box B13	2.8
60	115998	Hs.336629	AA448488	ESTs; Weakly similar to zinc finger prot	2.8
	112160	Hs.25067	R49116	EST	2.8
	120428	Hs.173694	AA236822	ESTs; Moderately similar to (defline not	2.8
	106241	Hs.6019	AA430108	ESTs	2.8
	131060	Hs.22564	AA160890	myosin VI	2.8
65	111383	Hs.40919	N94527	ESTs	2.8
	102123	Hs.1594	U14518	centromere protein A (17kD)	2.8
	102722	Hs.79981	U79242	Human clone 23560 mRNA sequence	2.8
	129887	Hs.274324	W92041	PCAF associated factor 65 alpha	2.8
	126663	Hs.181297	AA714635	ESTs	2.8

104367	Hs.134342	H17438	ESTs; Weakly similar to seventransmembra	2.8
107316	Hs.193700	T63174	ESTs; Moderately similar to !!!! ALU SUB	2.8
128059	Hs.145096	AA972446	ESTs	2.8
124447		N48000	ESTs	2.8
111398	Hs.125565	R00086	deafness; X-linked 1; progressive	2.8
134085	Hs.79018	U20979	chromatin assembly factor I (150 kDa)	2.8
124788	Hs.100912	R43543	ESTs	2.8
112248	Hs.326416	R51361	ESTs	2.8
121309	Hs.97312	AA402482	ESTs	2.8
103076	Hs.75319	X59618	ribonucleotide reductase M2 polypeptide	2.8
107071	Hs.35198	AA609053	ESTs	2.8
104425	Hs.35380	H88496	ESTs	2.8
132991	Hs.62245	AA446906	solute carrier family 25 (mitochondrial	2.8
104968	Hs.29669	AA084602	ESTs	2.8
121153	Hs.97694	AA399640	ESTs	2.8
131216	Hs.243901	D31058	ESTs	2.8
109682	Hs.22869	F09299	ESTs	2.8
131990	Hs.168818	H77734	ESTs; Moderately similar to roundabout 1	2.8
132027	Hs.181444	N78844	ESTs; Weakly similar to R12C12.6 [C.eleg	2.8
127383	Hs.190478	AA447990	ESTs	2.8
132598	Hs.530	M81379	collagen; type IV; alpha 3 (Goodpasture	2.8
101121	Hs.1313	L09753	tumor necrosis factor (ligand) superfami	2.8
123000	Hs.105640	AA479347	ESTs	2.8
121329	Hs.1755	AA404324	ESTs	2.8
100481	Hs.121489	HG1098-HT1098	Cystatin D	2.7
113803	Hs.283683	W42789	ESTs	2.7
110934	Hs.169001	N48708	ESTs; Weakly similar to cytochrome P-450	2.7
432888		T86823	ESTs	2.7
121802	Hs.188898	AA424328	ESTs	2.7
130396	Hs.155313	AB002331	Human mRNA for KIAA0333 gene; partial cd	2.7
121103	Hs.97697	AA398936	ESTs; Weakly similar to (defline not ava	2.7
131129	Hs.23240	R27296	ESTs	2.7
130943	Hs.272429	D50855	calcium-sensing receptor (hypocalciuric	2.7
134676	Hs.87819	W28051	ESTs; Weakly similar to keratin 9; cytos	2.7
111900	Hs.25318	R39044	ESTs	2.7
106025	Hs.173334	AA412063	ESTs	2.7
126144	Hs.40639	N39696	yx92a07.r1 Soares melanocyte 2NbHM Homo	2.7
103248	Hs.75262	X77383	cathepsin O	2.7
127230	Hs.274170	H30501	Homo sapiens Opa-interacting protein OIP	2.7
101584	Hs.84072	M35252	transmembrane 4 superfamily member 3	2.7
124131	Hs.167489	H19980	ESTs	2.7
129689	Hs.77873	AA130156	ESTs	2.7
132892	Hs.9973	W92797	ESTs	2.7
120827	Hs.132967	AA347717	ESTs	2.7
134579	Hs.85963	N23222	ESTs; Moderately similar to !!!! ALU SUB	2.7
106149	Hs.256301	AA424881	ESTs	2.7
132037	Hs.332541	AA203649	ESTs; Weakly similar to HEM45 [H.sapiens	2.7
130542	Hs.179825	U64675	Human sperm membrane protein BS-63 mRNA,	2.7
122851	Hs.99598	AA463627	ESTs	2.7
134983	Hs.196384	D28235	prostaglandin-endoperoxide synthase 2 (p	2.7
120537	Hs.160422	AA262790	ESTs	2.7
131036	Hs.174140	X64330	ATP citrate lyase	2.7
133889	Hs.211582	AA099391	ESTs	2.7
128847	Hs.106529	AA424199	zv81e01.r1 Soares_total_fetus_Nb2HF8_9w	2.7
112755	Hs.306044	R93802	ESTs	2.7
423239		AA323591	EST26392 Cerebellum II Homo sapiens cDNA	2.7
105031	Hs.12321	AA127240	ESTs	2.7
126021	Hs.187516	AA775894	ESTs	2.7
102116		U13706	Human ELAV-like neuronal protein 1 isofo	2.7
133394	Hs.237225	R16759	ESTs; Weakly similar to (defline not ava	2.7
104267	Hs.278439	C00358	ESTs	2.7
107614	Hs.40241	AA004878	ESTs; Highly similar to (defline not ava	2.7
129809	Hs.1259	X55283	asialoglycoprotein receptor 2	2.7
112109	Hs.283309	R45221	ESTs; Weakly similar to !!!! ALU SUBFAMI	2.7
128422		T85681	yd60c06.r1 Soares fetal liver spleen 1NF	2.7
109494	Hs.43899	AA233702	ESTs	2.7
118696	Hs.292284	N72086	Homo sapiens RNA polymerase III largest	2.7
106053	Hs.36727	AA416963	ESTs; Highly similar to histone H2A [H.s	2.7
104440	Hs.284380	L20492	gamma-glutamyltransferase 1	2.7

129426	Hs.111323	AA412087	EST; Highly similar to (defline not avai	2.7
123798		AA620411	small inducible cytokine A5 (RANTES)	2.7
106716	Hs.238928	AA464962	ESTs	2.7
103663		Z78291	Z78291 Homo sapiens brain fetus Homo sap	2.7
114162	Hs.22265	Z38909	ESTs	2.7
113063	Hs.5027	T32438	ESTs	2.7
127897		AA773857	af80c09.r1 Soares_NhHMPu_S1 Homo sapiens	2.7
130621	Hs.16803	AA621718	ESTs; Weakly similar to (defline not ava	2.7
116245	Hs.42796	AA479958	ESTs; Highly similar to (defline not ava	2.7
125499		R11878	yf49d11.r1 Soares infant brain 1NIB Homo	2.7
133960	Hs.77899	M19267	tropomyosin 1 (alpha)	2.7
104470	Hs.246358	N28843	ESTs; Weakly similar to Similar to colla	2.7
134982	Hs.92308	N46086	ESTs	2.7
106803	Hs.284295	AA479114	ESTs	2.7
104899	Hs.285574	AA054726	ESTs	2.7
125401	Hs.337585	AI204637	ESTs; Moderately similar to KIAA0350 [H.	2.7
111253	Hs.15768	N70042	ESTs; Moderately similar to !!!! ALU SUB	2.7
118449	Hs.164478	N66413	ESTs; Weakly similar to (defline not ava	2.7
134507	Hs.84318	M63488	replication protein A1 (70kD)	2.7
121609	Hs.98185	AA416867	EST	2.7
113835	Hs.27475	W56590	ESTs	2.7
113962	Hs.285290	W86375	ESTs; Highly similar to (defline not ava	2.7
121913	Hs.98558	AA428062	ESTs	2.7
108194	Hs.216717	AA057250	ESTs	2.7
130799	Hs.12696	AA464273	ESTs	2.7
123184	Hs.18166	AA489072	Homo sapiens mRNA for KIAA0870 protein;	2.7
103420	Hs.173497	X97065	SEC23-like protein B	2.7
106186	Hs.6315	AA427398	acetylserotonin N-methyltransferase-like	2.7
101349		L77559	Homo sapiens DGS-B partial mRNA	2.7
112954	Hs.6655	T16559	ESTs	2.7
133054	Hs.291079	R07876	ESTs; Weakly similar to unknown [S.cerev	2.7
128131	Hs.25640	AI283162	claudin 3	2.6
101864	Hs.75777	M95787	transgelin	2.6
111948	Hs.26303	R40752	ESTs	2.6
130145	Hs.151051	U07620	protein kinase mitogen-activated 10 (MAP	2.6
126507	Hs.23964	AI362218	ESTs	2.6
117903	Hs.47111	N50740	ESTs	2.6
116345	Hs.199067	AA496981	ESTs	2.6
132227	Hs.4248	AA412620	ESTs	2.6
125746	Hs.274256	H03574	yj42b06.r1 Soares placenta Nb2HP Homo sa	2.6
105073	Hs.89463	AA137034	ESTs	2.6
102764		U82310	Homo sapiens unknown protein mRNA, parti	2.6
131367	Hs.173933	AA456687	ESTs	2.6
130792	Hs.19500	AA307896	nuclear localization signal deleted in v	2.6
107427	Hs.46736	W26975	ESTs	2.6
117477	Hs.44175	N30328	ESTs	2.6
106290	Hs.16364	AA435542	ESTs	2.6
126829	Hs.7910	R11547	ESTs	2.6
118836	Hs.173001	N79820	ESTs	2.6
100147	Hs.136348	D13666	osteoblast specific factor 2 (fasciclin	2.6
104278	Hs.109253	C02582	ESTs; Highly similar to (defline not ava	2.6
135051	Hs.83484	C15324	ESTs	2.6
126081	Hs.227835	AI346024	collagen; type I; alpha 1	2.6
123579		AA608983	af5d4.s1 Soares_testis_NHT Homo sapiens	2.6
130115	Hs.149923	M31627	X-box binding protein 1	2.6
101434	Hs.1430	M20218	coagulation factor XI (plasma thrombopla	2.6
122962	Hs.104720	AA478429	ESTs; Moderately similar to !!!! ALU SUB	2.6
126151	Hs.40808	AA324743	ESTs	2.6
128925	Hs.21851	D61676	Homo sapiens mRNA; cDNA DKFZp586J2118 (f	2.6
128919	Hs.103391	L27559	insulin-like growth factor binding prote	2.6
130296	Hs.154103	R09286	LIM protein (similar to rat protein kina	2.6
128402	Hs.191637	AA457244	ESTs	2.6
129273	Hs.109968	W63783	ESTs	2.6
125483	Hs.7788	F07759	ESTs	2.6
132953	Hs.321264	AA029927	ESTs	2.6
130963	Hs.21639	U57099	nuclear protein; marker for differentiat	2.6
120614	Hs.194154	AA284281	ESTs; Weakly similar to !!!! ALU SUBFAMI	2.6
123251	Hs.103267	AA490858	ESTs; Moderately similar to Rabin3 [R.no	2.6
121710	Hs.96744	AA419011	ESTs	2.6

125428	Hs.851	W74608	ESTs; Highly similar to (defline not ava	2.6
115906	Hs.82302	AA436616	ESTs	2.6
108432		AA076626	Homo sapiens clone 23851 mRNA sequence	2.6
126191	Hs.191911	H97728	ESTs	2.6
106164	Hs.281434	AA425773	ESTs	2.6
111519	Hs.268615	R08165	ESTs	2.6
134590	Hs.173840	W58612	ESTs	2.6
102565		U59748	Human desert hedgehog (hDHH) mRNA, parti	2.6
129879	Hs.13109	AA194973	ESTs	2.6
114264	Hs.334609	Z40074	ESTs	2.6
106236	Hs.21104	AA429951	ESTs	2.6
135192	Hs.321709	AF000234	purinergic receptor P2X; ligand-gated io	2.6
109833	Hs.29889	H00580	ESTs	2.6
105756	Hs.8535	AA303088	ESTs; Weakly similar to transformation-r	2.6
121422	Hs.97967	AA406210	ESTs	2.6
130417	Hs.155485	U58522	Human huntingtin interacting protein (HI	2.6
124312	Hs.102329	H94647	ESTs	2.6
108998	Hs.97199	AA156058	ESTs	2.6
127081	Hs.180591	R88362	ESTs; Weakly similar to weak similarity	2.6
129574	Hs.11463	AA458603	ESTs; Weakly similar to (defline not ava	2.6
112410	Hs.26904	R61680	ESTs	2.6
123929	Hs.112981	AA621364	ESTs	2.6
122905	Hs.104835	AA470070	ESTs	2.6
116399	Hs.110637	AA599729	Homo sapiens homeobox protein A10 (HOXA1	2.6
130279	Hs.153934	AA424044	core-binding factor; runt domain; alpha	2.6
130021	Hs.1435	M24470	guanosine monophosphate reductase	2.6
100585	Hs.199160	HG2367-HT2463	Trithorax Homolog Hrx	2.6
104965	Hs.30177	AA084104	ESTs	2.6
117711	Hs.46485	N45201	EST	2.6
124792	Hs.48712	R44357	ESTs	2.6
111299	Hs.74313	N73808	ESTs	2.6
103616	Hs.32971	Z46973	phosphoinositide-3-kinase; class 3	2.6
133629	Hs.195614	D13642	KIAA0017 gene product	2.6
126484	Hs.169977	AI086782	ESTs	2.6
100858		HG4245-HT4515	Forkhead Family Afx1	2.6
133547	Hs.301927	X02883	T-cell receptor; alpha (V;D;J;C)	2.6
126680	Hs.133865	F07097	ESTs	2.6
125739	Hs.92137	AA426557	v-myc avian myelocytomatosis viral oncog	2.6
102276	Hs.10247	U30999	Human (memc) mRNA, 3'UTR	2.6
105586	Hs.191538	AA279137	ESTs	2.6
103978	Hs.34136	AA307443	ESTs	2.6
125054	Hs.268601	T80622	ESTs; Weakly similar to (defline not ava	2.6
114212	Hs.21201	Z39338	ESTs; Highly similar to (defline not ava	2.6
116959	Hs.40022	H79310	EST	2.6
109228	Hs.306995	AA193366	ESTs	2.6
133989	Hs.78202	U29175	SWI/SNF related; matrix associated; acti	2.6
100640	Hs.182183	HG2743-HT2845	Caldesmon 1, Alt. Splice 3, Non-Muscle	2.6
133093	Hs.285996	AA598749	ESTs	2.6
114306	Hs.6540	Z40861	ESTs	2.6
106060	Hs.171391	AA417287	C-terminal binding protein 2	2.5
107748	Hs.60772	AA017258	EST	2.5
100134	Hs.49	D13264	macrophage scavenger receptor 1	2.5
133969	Hs.78	U13044	GA-binding protein transcription factor;	2.5
130992	Hs.74316	AA455001	ESTs	2.5
127493	Hs.291701	AA808081	oc39a08.s1 NCI_CGAP_GCB1 Homo sapiens cD	2.5
132869	Hs.203961	N26855	ESTs	2.5
117570	Hs.44583	N34415	EST	2.5
124644	Hs.109654	N91279	ESTs	2.5
103558	Hs.2785	Z19574	keratin 17	2.5
132883	Hs.5897	AA047151	ESTs	2.5
102009	Hs.82643	U02680	protein tyrosine kinase 9	2.5
116058	Hs.20159	AA454156	ESTs	2.5
121989	Hs.193784	AA430044	ESTs	2.5
131257	Hs.24908	AA256042	ESTs	2.5
100320	Hs.75275	D50916	homolog of yeast (S. cerevisiae) ufd2	2.5
102959	Hs.121524	X15722	glutathione reductase	2.5
132969	Hs.6166	AA047616	ESTs	2.5
130869	Hs.2057	AA128100	uridine monophosphate synthetase (orotat	2.5
129645	Hs.118131	L38928	5;10-methenyltetrahydrofolate synthetase	2.5

126399	Hs.83883	AA128075	zl16d08.r1 Soares_pregnant_uterus_NbHPU	2.5
134069	Hs.78935	U29607	Homo sapiens eIF-2-associated p67 homolo	2.5
109816	Hs.61960	F11013	ESTs; Weakly similar to KIAA0176 [H.sapi	2.5
134801	Hs.89695	X02160	insulin receptor	2.5
104232	Hs.10587	AB002351	Human mRNA for KIAA0353 gene; partial cd	2.5
107361	Hs.159486	U72513	Human RPL13-2 pseudogene mRNA; complete	2.5
106057	Hs.289074	AA417067	ESTs	2.5
134252	Hs.80720	AA031782	Homo sapiens mRNA; cDNA DKFZp586B1722 (f	2.5
128062	Hs.105547	AA379500	ESTs	2.5
110009	Hs.6614	H10933	ESTs	2.5
111375	Hs.20432	N93696	ESTs	2.5
122642	Hs.99361	AA454186	ESTs	2.5
127999	Hs.69851	AA837495	ESTs; Weakly similar to Wiskott-Aldrich	2.5
105029	Hs.13268	AA126855	ESTs	2.5
105082	Hs.26765	AA143763	ESTs; Weakly similar to Similarity to S.	2.5

TABLE 1A show the accession numbers for those primekeys lacking unigeneID's for Table 1. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Pkey: Unique Eos probeset identifier number
CAT number: Gene cluster number
Accession: Genbank accession numbers

Pkey	CAT number	Accessions
108552	111555_1	AA071210 AA069899 AA071438 AA084912 AA084803 AA079371 AA079370
126023	1596090_1	H57661 H58881
126086	1606216_1	H75681 H70975
102565	32479_1	AB010994 U59748 AA064660
101964	48158_-7	S81578
125499	1562851_1	H10543 R11878
125596	1708455_1	R25698 R56582 R56018
118417	37186_1	AF080229 AF080231 AF080230 AF080232 AF080233 AF080234 BE550633 AI636743 AW614951 BE467547 AI680833 AI633818 N29986 U87592 U87593 U87590 U87591 S46404 U87587 AA463992 AW206802 AI970376 AI583718 AI672574 N25695 AW665466 AI818326 AA126128 AI480345 AW013927 AA248638 AI214968 AA204735 AA207155 AA206262 AA204833 AW003247 AW496808 AI080480 AI631703 AI651023 AI867418 AW818140 AA502500 AI206199 AI671282 AI352545 BE501030 AI652535 BE465762 AA206331 AW451866 AA471088 AA206342 AA204834 AA206100 AW021661 AA332922 N66048 AA703396 H92278 AW139734 H92683 U87589 U87595 H69001 U87594 BE466420 AI624817 BE466611 AI206344 AA574397 AA348354 AI493192
125661	327827_1	AA491830 R50173 R55192 R50320 AI732306 AI732305 AI820727 AI820728 R55191 R50319 R50227
125957	1583542_1	H41694 H45213
125982	1766315_1	R98091 W92898
127248	227560_1	AA364195 AA325029 AW962050
103731	112052_1	AA070545 AA131490 AA131373
127261	231687_1	AA330501 AA661567
127265	232391_1	AA331503 AA332751 AW962542
126659	1541209_1	T16245 R19694 F13545 H10299 T66048 T65279 H18006
127315	37938_1	AF116622 AI114507 AA640834 AA377999
103806	112618_1	AA130614 AA071410
128104	502608_1	AA906093 AA971000
104602	524482_2	H47610 R86920
128152	297868_1	F07973 R20353 AA442660
128422	1811283_1	T77794 T85681
127897	446527_1	AA773681 AA773857
106566	120358_1	BE298210 AI672315 AW086489 BE298417 AA455921 AA902537 BE327124 R14963 AA085210 AW274273 AI333584 AI369742 AI039658 AI885095 AI476470 AI287650 AI885299 AI985381 AW592624 AW340136 AI266556 AA456390 AI310815 AA484951
129735	44573_2	AI950087 N70208 R97040 N36809 AI308119 AW967677 N35320 AI251473 H59397 AW971573 R97278 W01059 AW967671 AA908598 AA251875 AI820501 AI820532 W87891 T85904 U71456 T82391 BE328571 T75102 R34725 AA884922 BE328517 AI219788 AA884444 N92578 F13493 AA927794 AI560251 AW874068 AL134043 AW235363 AA663345 AW008282 AA488964 AA283144 AI890387 AI950344 AI741346 AI689062 AA282915 AW102898 AI872193 AI763273 AW173586 AW150329 AI653832 AI762688 AA988777 AA488892 AI356394 AW103813 AI539642 AA642789 AA856975 AW505512 AI961530 AW629970 BE612881 AW276997 AW513601 AW512843 AA044209 AW856538 AA180009 AA337499 AW961101 AA251669 AA251874 AI819225 AW205862 AI683338 AI858509 AW276905 AI633006 AA972584 AA908741 AW072629 AW513996 AA293273 AA969759 N75628 N22388 H84729 H60052 T92487 AI022058 AA780419 AA551005 W80701 AW613456 AI373032 AI564269 F00531 H83488 W37181 W78802 R66056 AI002839 R67840 AA300207 AW959581 T63226 F04005
123147	219802_-2	AA487961
130529	158447_1	AA178953 AA192740
123579	genbank_AA608983	AA608983
109175	genbank_AA180496	AA180496
100789	tigr_HT4163	S67998
100858	tigr_HT4515	U10072

123798	579959_1	AA620411	AA287491
102116	entrez_U13706	U13706	
102398	entrez_U42359	U42359	
102764	entrez_U82310	U82310	
118475	genbank_N66845	N66845	
104776	genbank_AA026349	AA026349	
104787	genbank_AA027317	AA027317	
113702	genbank_T97307	T97307	
113938	genbank_W81598	W81598	
122635	genbank_AA454085	AA454085	
108407	genbank_AA075519	AA075519	
108432	genbank_AA076626	AA076626	
108555	genbank_AA084963	AA084963	
101349	entrez_L77559	L77559	
124447	genbank_N48000	N48000	
119071	genbank_R31180	R31180	
103520	entrez_Y10511	Y10511	
103663	genbank_Z78291	Z78291	
128046	877605_1	AA873285	AI025762
126959	546044_1	AA199853	AA206355
123465	genbank_AA599033	AA599033	

MISSING AT THE TIME OF PUBLICATION

TABLE 2: shows a preferred subset of the Accession numbers for genes found in Table 1 which are differentially expressed in prostate tumor tissue compared to normal prostate tissue.

Pkey:	Unique Eos probeset identifier number
ExAccn:	Exemplar Accession number, Genbank accession number
UnigeneID:	Unigene number
Unigene Title:	Unigene gene title
R1:	Ratio of tumor to normal body tissue (Relaxed ratio (87/70)

Pkey	ExAccn	UnigeneID	Unigene Title	R1
131919	AA121266	Hs.272458	ESTs	37.2
120328	AA196979	Hs.290905	ESTs; Weakly similar to (defline not ava	32.6
101486	M24902	Hs.1852	acid phosphatase; prostate	25.2
119073	R32894	Hs.279477	ESTs	24.8
133428	M34376	Hs.183752	microseminoprotein; beta-	23.8
128180	AA595348	Hs.171995	kallikrein 3; (prostate specific antigen	21.4
104080	AA402971	Hs.57771	Homo sapiens mRNA for serine protease (T	18.9
127537	AA569531	Hs.162859	ESTs	18.6
131665	R22139	Hs.30343	ESTs	17.4
101050	K01911	Hs.1832	neuropeptide Y	17.3
130771	N48056	Hs.1915	folate hydrolase (prostate-specific memb	17
107485	W63793	Hs.262476	S-adenosylmethionine decarboxylase 1	16.7
106155	AA425309	Hs.33287	ESTs	16.5
129534	R73640	Hs.11260	ESTs	16.4
100569	HG2261-HT2351		Antigen, Prostate Specific, Alt. Splice	16
101889	S39329	Hs.181350	kallikrein 2; prostatic	15.4
135389	U05237	Hs.99872	fetal Alzheimer antigen	15
133944	AA045870	Hs.7780	ESTs	12.5
130974	X57985	Hs.2178	H2B histone family; member Q	11.8
114768	AA149007	Hs.182339	ESTs	11.8
104660	AA007160	Hs.14846	ESTs	11.4
131061	N64328	Hs.268744	ESTs; Moderately similar to KIAA0273 [H.	10.9
126645	AI167942	Hs.61635	Homo sapiens BAC clone RG041D11 from 7q2	10.7
135153	N40141	Hs.95420	Homo sapiens mRNA for JM27 protein; comp	10.6
107033	AA599629	Hs.113314	ESTs	10.6
118417	N66048		ESTs; Weakly similar to polymerase [H.sa	10.5
126758	W37145	Hs.293960	ESTs	10.2
107102	AA609723	Hs.30652	ESTs	10.1
116787	H28581	Hs.15641	ESTs	10.1
115719	AA416997	Hs.59622	ESTs	10
123209	AA489711	Hs.203270	ESTs	9.9
101664	M60752	Hs.121017	H2A histone family; member A	9.8
112971	T17185	Hs.83883	ESTs	9.7
117984	N51919	Hs.106778	ESTs	9.7
129523	M30894	Hs.274509	T-cell receptor; gamma cluster	9.4
132964	AA031360	Hs.167133	ESTs	9.2
121853	AA425887	Hs.98502	ESTs	9
119617	W47380	Hs.55999	ESTs	8.9
105627	AA281245	Hs.23317	ESTs	8.8
101461	M22430	Hs.76422	phospholipase A2; group IIA (platelets;	8.7
124526	N62096	Hs.293185	yz61c5.s1 Soares_multiple_sclerosis_2NbH	8.5
133845	T68510	Hs.76704	ESTs	8.2
133354	AA055552	Hs.334762	ESTs; Weakly similar to KIAA0319 [H.sapi	8.1
119018	N95796	Hs.278695	ESTs	8
100394	D84276	Hs.66052	CD38 antigen (p45)	8
106579	AA456135	Hs.23023	ESTs	7.6
114965	AA250737	Hs.72472	ESTs	7.4
112033	R43162	Hs.22627	ESTs	7.1
102398	U42359		Human N33 protein form 1 (N33) gene, exo	7
101201	L22524	Hs.2256	matrix metalloproteinase 7 (matrilysin;	6.9
101803	M86546	Hs.155691	pre-B-cell leukemia transcription factor	6.8
120562	AA280036	Hs.302267	ESTs; Weakly similar to W01A6.c [C.elega	6.8

109112	AA169379	Hs.257924	ESTs	6.8
109795	F10707	Hs.326416	ESTs	6.7
130336	X07730	Hs.171995	kallikrein 3; (prostate specific antigen	6.6
131425	AA219134	Hs.26691	ESTs	6.6
132802	AA490969	Hs.59838	ESTs	6.6
133724	U07919	Hs.75746	aldehyde dehydrogenase 6	6.5
120215	Z41050	Hs.108787	Homo sapiens Mcd4p homolog mRNA; complet	6.5
131881	AA010163	Hs.3383	upstream regulatory element binding prot	6.5
100727	X07290	Hs.334786	Human HF.12 gene mRNA	6.3
121770	AA421714	Hs.278428	Homo sapiens mRNA for KIAA0896 protein;	6.3
123475	AA599267	Hs.250528	ESTs; Weakly similar to ANKYRIN; BRAIN V	6.3
133061	AB000584	Hs.296638	prostate differentiation factor	6.3
116429	AA609710	Hs.279923	ESTs; Weakly similar to similar to GTP-b	6.2
101233	L29008	Hs.878	sorbitol dehydrogenase	6.2
104691	AA011176	Hs.37744	ESTs	6.2
127248	AA325029		EST27953 Cerebellum II Homo sapiens cDNA	6.2
105500	AA256485	Hs.222399	ESTs	6.1
130828	AA053400	Hs.203213	ESTs	5.9
115357	AA281793	Hs.72988	ESTs	5.8
116334	AA491457	Hs.48948	ESTs	5.7
120132	Z38839	Hs.125019	ESTs; Weakly similar to !!!! ALU SUBFAMI	5.6
106375	AA443993	Hs.289072	ESTs	5.6
124777	R41933	Hs.140237	ESTs; Weakly similar to neuronal thread	5.6
101791	M83822	Hs.62354	Human beige-like protein (BGL) mRNA; par	5.5
117698	N41002	Hs.45107	ESTs	5.5
122041	AA431407	Hs.98732	Homo sapiens Chromosome 16 BAC clone CIT	5.5
133723	AA088851	Hs.262476	S-adenosylmethionine decarboxylase 1	5.5
113938	W81598		ESTs	5.4
133015	AA047036	Hs.246315	ESTs	5.4
108186	AA056482	Hs.7760	ESTs	5.3
104466	N25110	Hs.326392	Human guanine nucleotide exchange factor	5.3
104033	AA365031	Hs.98944	ESTs	5.3
110844	N31952	Hs.167531	ESTs; Weakly similar to (defline not ava	5.3
129056	H70627	Hs.108336	ESTs; Weakly similar to !!!! ALU SUBFAMI	5.3
133493	AA284143	Hs.194369	Homo sapiens chromosome 1 atrophin-1 rel	5.3
129184	W26769	Hs.109201	ESTs; Highly similar to (defline not ava	5.2
101448	M21389	Hs.195850	keratin 5 (epidermolysis bullosa simplex	5.1
116188	AA464728	Hs.184598	ESTs; Weakly similar to !!!! ALU SUBFAMI	5.1
105921	AA402613	Hs.169119	ESTs	5.1
103375	X91868	Hs.54416	sine oculis homeobox (Drosophila) homolo	5.1
128871	AA400271	Hs.106778	ESTs; Highly similar to (defline not ava	5.1
116238	AA479362	Hs.47144	ESTs	5
102913	X07696	Hs.80342	keratin 15	5
103011	X52541	Hs.326035	early growth response 1	5
118981	N93839	Hs.39288	ESTs; Weakly similar to !!!! ALU SUBFAMI	5

TABLE 2A shows the accession numbers for those primekeys lacking unigeneID's for Table 2. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Pkey: Unique Eos probeset identifier number
CAT number: Gene cluster number
Accession: Genbank accession numbers

Pkey	CAT number	Accession
118417	37186_1	AF080229 AF080231 AF080230 AF080232 AF080233 AF080234 BE550633 AI636743 AW614951 BE467547 AI680833 AI633818 N29986 U87592 U87593 U87590 U87591 S46404 U87587 AA463992 AW206802 AI970376 AI583718 AI672574 N25695 AW665466 AI818326 AA126128 AI480345 AW013827 AA248638 AI214968 AA204735 AA207155 AA206262 AA204833 AW003247 AW496808 AI080480 AI631703 AI651023 AI867418 AW818140 AA502500 AI206199 AI671282 AI352545 BE501030 AI652535 BE465762 AA206331 AW451866 AA471088 AA206342 AA204834 AA206100 AW021661 AA332922 N66048 AA703396 H92278 AW139734 H92683 U87589 U87595 H69001 U87594 BE466420 AI624817 BE466611 AI206344 AA574397 AA348354 AI493192
127248	227560_1	AA364195 AA325029 AW962050
107033	235652_1	AI141999 AA730176 R44544 R41778 AW300793 AW966157 AA918501 AA599629 AI082195 AI196537 AW006520 AW236663 AW151420 AI826987 AI810832 AI669102 AI201981 N27331 AA335566 T84622 BE085347 BE085269
102398	entrez_U42359	U42359
113938	genbank_W81598	W81598

TABLE 3: shows genes, including expression sequence tags, differentially expressed in prostate tumor tissue compared to normal tissue as analyzed using the Affymetrix/Eos Hu02 GeneChip array. Shown are the relative amounts of each gene expressed in prostate tumor samples and various normal tissue samples showing the highest expression of the gene.

Pkey: *Unique Eos probeset identifier number*
ExAccn: Exemplar Accession number, Genbank accession number
UnigeneID: Unigene number
Unigene Title: Unigene gene title
R1: Ratio of tumor to normal body tissue

Pkey	ExAccn	UnigeneID	Unigene Title	R1
100131	D12485	Hs.11951	phosphodiesterase I/nucleotide pyrophosp	6.3
100235	D29954	Hs.13421	KIAA0056 protein	5.1
100570	HG2261-HT2352		Hs.171995	Antigen, Prostate Specific, Alt. Splice 9
100819	HG4020-HT4290		Hs.2387	Transglutaminase 10.5
101063	L00354	Hs.80247	*cholecystokinin*	8.5
101247	L33801	Hs.78802	glycogen synthase kinase 3 beta	4.7
101416	M17254	Hs.279477	v-ets avian erythroblastosis virus E26 o	4.7
101447	M21305		Human alpha satellite and satellite 3 ju	11
101485	M24736	Hs.89546	selectin E (endothelial adhesion molecul	9.8
101514	M28214	Hs.123072	RAB3B; member RAS oncogene family	6.2
101626	M57399	Hs.44	pleiotrophin (heparin binding growth fac	8.4
101663	M60750	Hs.2178	H2B histone family; member A	4.9
101758	M77836	Hs.79217	pyrroline-5-carboxylate reductase 1	5.4
101768	M81118	Hs.78989		7.5
101817	M88163	Hs.152292	SWI/SNF related; matrix associated; acti	5.5
101888	M99701	Hs.95243	transcription elongation factor A (SII)-	5.7
102031	U04898	Hs.2156	RAR-related orphan receptor A	13.2
102052	U07559	Hs.505	ISL1 transcription factor; LIM/homeodoma	8.9
102221	U24576	Hs.3844	LIM domain only 4	5.6
102233	U26173	Hs.79334	*nuclear factor; interleukin 3 regulated*	7.4
102302	U33052	Hs.69171	protein kinase C-like 2	8.2
102348	U37519	Hs.87539	aldehyde dehydrogenase 8	5.9
102457	U48807	Hs.2359	dual specificity phosphatase 4	5.1
102473	U49957	Hs.180398	LIM domain-containing preferred transloc	5.7
102699	U71207	Hs.29279	eyes absent (Drosophila) homolog 2	9
102698	U75272	Hs.1867	progastricsin (pepsinogen C)	10.6
102751	U80034	Hs.68583	mitochondrial intermediate peptidase	15.6
102823	U90914	Hs.5057	carboxypeptidase D	4.9
102869	X02544	Hs.572	orosomucoid 1	22.6
103031	X54667	Hs.123114	cystatin S	4.7
103043	X55733	Hs.93379	eukaryotic translation initiation factor	4.9
103093	X60708	Hs.44926	dipeptidylpeptidase IV (CD26; adenosine	5.8
103376	X92098	Hs.323378	*coated vesicle membrane protein*	5.2
103401	X95240	Hs.54431	specific granule protein (28 kDa); cyste	7.4
103613	Z46629	Hs.2316	SRY (*sex-determining region Y*)-box 9 (ca	5.2
103677	Z63806		H.sapiens mRNA for axonemal dynein heavy	4.9
103962	AA298180	Hs.83243	ESTs	6
104084	AA410529	Hs.30732	ESTs	6.4
104257	AF006265	Hs.9222	estrogen receptor-binding fragment-assoc	6.8
104301	D45332	Hs.6783	ESTs	10.5
104769	AA025887	Hs.293943	ESTs; Weakly similar to !!!! ALU SUBFAMI	6.3
104851	AA040882	Hs.10290	U5 snRNP-specific 40 kDa protein (hPrp8-	4.9
104896	AA054228	Hs.23165	ESTs	5.8
104956	AA074880	Hs.20509	ESTs; Weakly similar to hypothetical pro	6.4
104957	AA074919	Hs.10026	ESTs; Weakly similar to ORF YJL063c [S.c	4.8
104967	AA084506	Hs.291000	ESTs	6.5
105099	AA150776	Hs.23729	Homo sapiens clone 24405 mRNA sequence	7
105298	AA233459	Hs.26369	ESTs	5.1

105304	AA233553	Hs.190325	ESTs	4.7
105370	AA236476	Hs.22791	ESTs; Weakly similar to transmembrane pr	10.3
105427	AA251330	Hs.28248	ESTs	5
105542	AA261858	Hs.266957	ESTs; Weakly similar to heat shock prote	8.8
105628	AA281251	Hs.79828	ESTs; Weakly similar to putative zinc fi	5.5
105640	AA281623	Hs.6685	ESTs; Weakly similar to KIAA0742 protein	8
105645	AA282138	Hs.11325	ESTs	14
105691	AA287097	Hs.289068	transcription factor 4	6.3
105730	AA292701	Hs.5364	DKFZP564I052 protein	4.9
105808	AA393808	Hs.286131	KIAA0438 gene product	7
105826	AA398243	Hs.194477	ESTs; Moderately similar to similar to N	5
105903	AA401433	Hs.200016	ESTs; Weakly similar to diphosphoinosito	9.9
105906	AA401633	Hs.22380	ESTs	11.5
106065	AA417558	Hs.25206	ESTs	5.1
106094	AA419461	Hs.23317	ESTs	10.9
106157	AA425367	Hs.34892	ESTs	6.6
106184	AA426643	Hs.10762	ESTs	8.5
106211	AA428240	Hs.126083	ESTs	8.4
106213	AA428258	Hs.8769	Homo sapiens mRNA; cDNA DKFZp564E153 (fr	5.7
106272	AA432074	Hs.323099	ESTs	5.8
106369	AA443828	Hs.288856	ESTs	6.3
106400	AA447621	Hs.94109	ESTs	5.4
106474	AA450212	Hs.42484	Homo sapiens mRNA; cDNA DKFZp564C053 (fr	9.2
106507	AA452584	Hs.267819	protein phosphatase 1; regulatory (inhib	5.6
106523	AA453441	Hs.31511	ESTs	4.7
106532	AA453628	Hs.37443	ESTs	4.7
106557	AA455087	Hs.22247	ESTs	5.7
106575	AA456039	Hs.105421	ESTs	7.2
106618	AA459249	Hs.8715	ESTs; Weakly similar to Similarity with	5.6
106820	AA481037	Hs.12592	ESTs	5.4
106846	AA485223	Hs.34892	ESTs	5.3
106973	AA505141	Hs.11923	Human DNA sequence from clone 167A19 on	7.5
107110	AA609952	Hs.12784	KIAA0293 protein	6.1
107127	AA620504	Hs.179898	ESTs	7.1
107159	AA621340	Hs.10600	ESTs; Weakly similar to ORF YKR081c [S.c	5.2
107217	D51095	Hs.35861	DKFZP586E1621 protein	15.1
107365	U78294	Hs.111256	arachidonate 15-lipoxygenase; second typ	4.7
107630	AA007218	Hs.60178	ESTs	5.3
107734	AA016225	Hs.7517	ESTs	4.8
107760	AA018042	Hs.252085	EST	7.6
107997	AA037388	Hs.82223	Human DNA sequence from clone 141H5 on c	10.5
108012	AA039616	Hs.173334	ESTs	6.5
108520	AA084138	Hs.46786	ESTs	7.9
108583	AA088276	Hs.68826	ESTs	5.6
108613	AA100967	Hs.69165	ESTs	6
108664	AA113349	Hs.69588	EST	6.3
108677	AA115629	Hs.118531	ESTs	5.9
108807	AA129968	Hs.49376	ESTs; Weakly similar to PROTEIN PHOSPHAT	5.8
108910	AA136590		ESTs	5
108933	AA147224	Hs.337232	ESTs	12.7
108948	AA149579	Hs.118258	ESTs	6.8
109014	AA156790	Hs.262036	ESTs	15.3
109124	AA171529	Hs.183887	ESTs	6.1
109142	AA176438	Hs.41295	ESTs	5.1
109277	AA196332	Hs.86043	ESTs	5.5
109342	AA213620		Homo sapiens mRNA; cDNA DKFZp586M1418 (f6	
109562	F01811	Hs.187931	ESTs; Moderately similar to voltage-gate	10.8
109565	F01930	Hs.23648	ESTs	7
109648	F04600	Hs.7154	ESTs	9.9
109799	F10770	Hs.180378	Homo sapiens clone 669 unknown mRNA; com	6.4
109859	H02308	Hs.20792	ESTs	5.3
110181	H20276	Hs.31742	ESTs	16.8
110854	N32919	Hs.27931	ESTs	10
110924	N47938	Hs.12940	yy84a09.s1 Soares_multiple_sclerosis_2Nb	5.6
111046	N55514	Hs.318584	ESTs	6.9
111091	N59858	Hs.33032	Homo sapiens mRNA; cDNA DKFZp434N185 (fr	5.2
111157	N66613	Hs.99364	ESTs	5
111164	N66857	Hs.122489	ESTs; Weakly similar to !!!! ALU CLASS C	5.6
111221	N68869	Hs.15119	ESTs	6.2

111348	N90041	Hs.9585	ESTs	5.4
111353	N90430	Hs.6616	ESTs	5.3
111495	R07210	Hs.9683	ESTs	5.8
111540	R08850	Hs.9786	ESTs	6
111579	R10657	Hs.167115	KIAA0830 protein	12.6
111581	R10684	Hs.5794	ESTs	7.1
111734	R25375	Hs.128749	ESTs	6.2
111861	R37460	Hs.25231	ESTs	9.4
111870	R37778	Hs.18685	ESTs; Weakly similar to hypothetical pro	6.5
111937	R40431	Hs.14846	Homo sapiens mRNA; cDNA DKFZp564D016 (fr	4.8
111987	R42036	Hs.6763	KIAA0942 protein	6.4
112184	R49173	Hs.330242	ESTs	5.6
112286	R53765	Hs.158135	KIAA0981 protein	9.3
112380	R59740	Hs.5740	ESTs	4.7
112452	R63841	Hs.157461	ESTs	6
112601	R79111	Hs.78225	annexin A1	5.4
112753	R93696	Hs.169882	ESTs	5.8
112902	T09262	Hs.129190	ESTs	5.1
112984	T23457	Hs.289014	ESTs	4.9
113021	T23855	Hs.129836	KIAA1028 protein	10.8
113083	T40530	Hs.266957	ESTs; Weakly similar to heat shock prote	5.7
113200	T57773	Hs.10263	ESTs	7.3
113494	T88878	Hs.86538	ESTs	8.7
113849	W60439	Hs.8858	ESTs; Moderately similar to cbp146 [M.mu	4.9
113883	W72382	Hs.11958	oxidative 3 alpha hydroxysteroid dehydro	4.7
113950	W85765	Hs.30504	Homo sapiens mRNA; cDNA DKFZp434E082 (fr	6.7
113986	W87462	Hs.21894	ESTs	5.9
113989	W87544	Hs.268828	ESTs	4.7
114124	Z38595	Hs.125019	ESTs; Highly similar to KIAA0886 protein	21.3
114340	Z41395	Hs.143611	ESTs	9.6
114346	Z41450	Hs.130489	ESTs	5.2
114435	AA018216	Hs.164975	Bicaudal D (Drosophila) homolog 1	7.4
114463	AA025370	Hs.40109	KIAA0872 protein	8.2
114652	AA101416	Hs.107149	ESTs; Weakly similar to PTB-ASSOCIATED S	5.4
114721	AA131450	Hs.103822	ESTs	4.8
114730	AA133527	Hs.331328	ESTs; Weakly similar to The KIAA0138 gen	5.1
114833	AA234362	Hs.87159	ESTs; Moderately similar to CGI-66 prote	5.5
114860	AA235112	Hs.42179	ESTs; Moderately similar to similar to m	6.3
114884	AA235811	Hs.293672	ESTs	5.2
114895	AA236177	Hs.76591	KIAA0887 protein	4.7
114908	AA236545	Hs.54973	ESTs	5.2
114932	AA242751	Hs.16218	KIAA0903 protein	5.7
115084	AA255566	Hs.42484	Homo sapiens mRNA; cDNA DKFZp564C053 (fr	5.2
115140	AA258030	Hs.279938	ESTs; Weakly similar to supported by GEN	5.9
115468	AA287061	Hs.48499	ESTs; Highly similar to Bdelght protein	4.7
115583	AA398913	Hs.45231	LDOC1 protein	7.6
115709	AA412519	Hs.58279	ESTs	4.8
115772	AA423972	Hs.131740	ESTs	5
115774	AA424029	Hs.288390	ESTs; Moderately similar to dynamin; int	5.4
115776	AA424038	Hs.81897	ESTs	5
115821	AA427528	Hs.130965	ESTs; Weakly similar to ZINC FINGER PROT	13.7
115955	AA446121	Hs.44198	Homo sapiens BAC clone RG054D04 from 7q3	10.6
116024	AA451748	Hs.83883	Human DNA sequence from clone 718J7 on c	6.8
116108	AA457566	Hs.28777	ESTs	6
116117	AA459117	Hs.31575	SEC63; endoplasmic reticulum translocon	7.3
116146	AA460701	Hs.15423	ESTs	5.5
116296	AA489033	Hs.62601	Homo sapiens mRNA; cDNA DKFZp586K1318 (f	5.7
116379	AA521472	Hs.71252	ESTs	5.9
116393	AA599463	Hs.306051	protein phosphatase 2 (formerly 2A); reg	5.9
116401	AA599963	Hs.59698	ESTs	7.9
116416	AA609219	Hs.39982	ESTs	9.2
116587	D59325	Hs.121429	ESTs	5.2
116601	D80055	Hs.45140	ESTs	4.9
116684	F09156	Hs.66095	ESTs	7.2
116722	F13654		HSFIH32 Stratagene cat#937212 (1992) Hom	5.5
116766	H13260	Hs.95097	ESTs	5.9
117453	N29568	Hs.108319	thyroid hormone receptor-associated prot	6.9
117557	N33920	Hs.44532	diubiquitin	4.8
117708	N45114	Hs.126280	ESTs	6.3

118001	N52151	Hs.47447	ESTs	11.4
118229	N62339	Hs.166254	heat shock 90kD protein 1; alpha	6.2
118599	N69207	Hs.203697	ESTs	5.8
118645	N70358	Hs.125180	growth hormone receptor	7.1
118873	N89861	Hs.44577	ESTs	6
118985	N94303	Hs.55028	ESTs	9.3
119107	R42424	Hs.63841	ESTs	6
119126	R45175	Hs.117183	ESTs	17.9
119271	T16387	Hs.65328	ESTs	6
119367	T78324	Hs.250895	ESTs	5
119721	W69440	Hs.48376	ESTs	15.4
119741	W70205	Hs.43670	kinesin family member 3A	10.1
119780	W72967	Hs.191381	ESTs; Weakly similar to hypothetical pro	5.3
120217	Z41078	Hs.66035	ESTs	4.8
120266	AA173939	Hs.205442	ESTs; Weakly similar to inner centromere	8.8
120294	AA190888	Hs.153881	ESTs; Highly similar to NY-REN-62 antige	4.9
120418	AA236010	Hs.26613	Homo sapiens mRNA; cDNA DKFZp586F1323 (f	4.7
120486	AA253400	Hs.137569	tumor protein 63 kDa with strong homolog	5.6
120524	AA261852	Hs.192905	ESTs	4.9
120571	AA280738	Hs.34892	ESTs	8.8
120596	AA282074	Hs.237323	ESTs	6.2
120713	AA292655	Hs.96557	ESTs	9.9
120992	AA398246	Hs.97594	ESTs	16.4
121429	AA406293	Hs.41167	ESTs	6.9
121503	AA412049	Hs.290347	ESTs	7.6
121512	AA412105	Hs.193736	ESTs	5.8
121816	AA424814	Hs.48827	ESTs	4.6
122027	AA431302	Hs.98721	EST; Weakly similar to N-copine [H.sapie	5.6
122294	AA437311	Hs.98927	ESTs	5.7
122411	AA446859	Hs.99083	ESTs	6.5
122791	AA460158	Hs.129836	KIAA1028 protein	12.4
122792	AA460225	Hs.99519	ESTs	5.1
122969	AA478539	Hs.104336	ESTs	4.9
123095	AA485724	Hs.27413	ESTs	5.4
123100	AA485957	Hs.306219	Homo sapiens clone 25032 mRNA sequence	5
123295	AA495981	Hs.250830	ESTs	4.7
123311	AA496252	Hs.105069	ESTs	7.4
123583	AA609006	Hs.111240	ESTs	9.1
123619	AA609200		ESTs	4.7
123645	AA609310	Hs.188691	ESTs	4.8
123709	AA609651	Hs.112742	ESTs	7
123958	C14333	Hs.108327	damage-specific DNA binding protein 1 (1	5
124178	H45996	Hs.97101	putative G protein-coupled receptor	6.8
124352	N21626	Hs.102406	ESTs	10.2
124357	N22401		yw37g07.s1 Morton Fetal Cochlea Homo sap	10.6
124515	N58172	Hs.109370	ESTs	14.2
124911	R88992	Hs.174195	ESTs	4.8
125154	W38419		ESTs	4.7
125992	W01626		za36e07.r1 Soares fetal liver spleen 1NF	5.1
126802	AA947601	Hs.97056	ESTs	5.1
126812	Z36290	Hs.173933	ESTs; Weakly similar to NUCLEAR FACTOR 1	4.6
127080	AA662913	Hs.190173	ESTs	5
127308	AA507628	Hs.334390	ESTs	4.8
127370	AI024352	Hs.70337	Immunoglobulin superfamily; member 4	4.7
127386	AI457411	Hs.106728	ESTs	4.8
127965	AA828760	Hs.292059	ESTs	4.8
128172	AI400862	Hs.265130	ESTs	5
128305	AI039722	Hs.279009	ESTs	5.8
128420	AI088155	Hs.41296	ESTs; Weakly similar to unknown [H.sapie	17
128467	AA176446	Hs.180428	ESTs; Weakly similar to hypothetical 43.	4.8
128610	L38608	Hs.10247	activated leucocyte cell adhesion molecu	7.9
128625	AA242816	Hs.102652	ESTs; Weakly similar to KIAA0437 [H.sapi	8.1
128651	AA446990	Hs.103135	ESTs	6.5
129088	AA215971	Hs.194431	KIAA0992 protein	5.2
129136	N26391	Hs.250723	ESTs	5.1
129171	AA234048	Hs.7753	calumenin	5.8
129229	AA211941	Hs.109643	polyadenylate binding protein-interactin	5.8
129386	N27524	Hs.260024	Cdc42 effactor protein 3	5.2
129467	AA410311	Hs.44208	ESTs	5.1

	129564	H22136	Hs.75295	guanylate cyclase 1; soluble; alpha 3	16.3
	129699	AA458578	Hs.12017	KIAA0439 protein; homolog of yeast ubiqu	9.2
	129821	F11019	Hs.12696	cortactin SH3 domain-binding protein	8.6
	129823	X00948	Hs.105314	relaxin 2 (H2)	9.1
5	129847	W46767	Hs.296178	ESTs; Weakly similar to RNA POLYMERASE I	5.4
	129912	AA047344	Hs.107213	ESTs; Highly similar to NY-REN-6 antigen	6.5
	129958	L20591	Hs.1378	annexin A3	5.1
	129977	J04076	Hs.1395	early growth response 2 (Krox-20 (Drosop	8.6
	130061	U82256	Hs.172851	arginase; type II	7.4
10	130241	U76313	Hs.153203	MyoD family inhibitor	4.9
	130466	N21679	Hs.180059	ESTs	5.8
	130541	X05608	Hs.211584	neurofilament; light polypeptide (68kD)	6.7
	130619	AA477739	Hs.12532	ESTs	6.4
	130925	N71935	Hs.169378	multiple PDZ domain protein	7.9
15	130938	AA013250	Hs.21398	ESTs; Moderately similar to PUTATIVE GLU	6.2
	130971	H20332	Hs.301444	signal sequence receptor; gamma (transla	6.4
	131066	F09006	Hs.22588	ESTs	5
	131126	F09012	Hs.181326	myotubularin related protein 2	6.4
	131310	J02960	Hs.2551	adrenergic; beta-2-; receptor; surface	7.9
20	131487	AA253220	Hs.27373	Homo sapiens mRNA; cDNA DKFZp564O1763 (f	5.9
	131561	X59841	Hs.294101	pre-B-cell leukemia transcription factor	7.6
	131562	U90551	Hs.28777	H2A histone family; member L	5.1
	131579	N62922	Hs.29088	ESTs	11
	131629	AA442119	Hs.238809	ESTs	4.9
25	131682	AA428368	Hs.30654	ESTs	4.8
	131699	R68657	Hs.90421	ESTs; Moderately similar to !!!! ALU SUB	6.5
	131795	N32724	Hs.32317	Sox-like transcriptional factor	5.6
	132053	H93381	Hs.38085	ESTs; Weakly similar to putative glycine	7.2
	132122	U65092	Hs.40403	Cbp/p300-interacting transactivator; wit	5.6
30	132191	AA449431	Hs.288361	KIAA0741 gene product	8
	132256	AA608856	Hs.431	murine leukemia viral (bmi-1) oncogene h	5.5
	132482	AA429478	Hs.238126	ESTs; Highly similar to CGI-49 protein [	6.6
	132533	AA021608	Hs.172510	ESTs	5.8
	132572	AA448297	Hs.237825	signal recognition particle 72kD	6.2
35	132581	R42266	Hs.52256	ESTs; Weakly similar to beta-TrCP protei	16
	132700	N47109	Hs.5521	ESTs	6.8
	132701	AA279359	Hs.55220	BCL2-associated athanogene 2	5.3
	132725	L41887	Hs.184167	splicing factor; arginine/serine-rich 7	7.8
	132783	N74897	Hs.278894	DEAD/H (Asp-Glu-Ala-Asp/His) box polypep	5.9
40	132790	X75535	Hs.168670	peroxisomal farnesylated protein	8
	132939	U76189	Hs.61152	exostoses (multiple)-like 2	5.2
	133142	F03321	Hs.65874	ESTs	5.2
	133342	U29589	Hs.7138	cholinergic receptor; muscarinic 3	10.3
	133434	AA278852	Hs.30212	ESTs	5.8
45	133453	M68941	Hs.73826	protein tyrosine phosphatase; non-recept	4.9
	133520	X74331	Hs.74519	primase; polypeptide 2A (58kD)	13.1
	133544	T33873	Hs.74624	protein tyrosine phosphatase; receptor t	4.6
	133608	D13315	Hs.75207	glyoxalase I	4.8
	133626	H75939	Hs.75277	Homo sapiens mRNA; cDNA DKFZp586M141 (fr	5
50	133633	D21262	Hs.75337	nucleolar phosphoprotein p130	6.3
	133797	S66431	Hs.76272	retinoblastoma-binding protein 2	6
	133928	N34096	Hs.7766	ubiquitin-conjugating enzyme E2E 1 (homo	5.4
	134095	U47414	Hs.79069	cyclin G2	5.2
	134249	N89827	Hs.80667	RALBP1 associated Eps domain containing	6.5
55	134321	AA418230	Hs.8172	ESTs	7
	134453	X70683	Hs.83484	SRY (sex determining region Y)-box 4	4.7
	134542	X57025	Hs.85112	insulin-like growth factor 1 (somatomedi	7.7
	134570	U66615	Hs.172280	SWI/SNF related; matrix associated; acti	6.4
	134592	U82613	Hs.289104	Alu-binding protein with zinc finger dom	5.4
60	134654	W23625	Hs.8739	ESTs; Weakly similar to ORF YGR200c [S.c	5
	134666	AA482319	Hs.8752	putative type II membrane protein	5.4
	134806	Z49099	Hs.89718	spermine synthase	6.7
	134951	AA431480	Hs.169358	ESTs	9.8
	135066	X04602	Hs.93913	interleukin 6 (interferon; beta 2)	5.7
65	135155	AA358268	Hs.166556	ESTs; Moderately similar to transcriptio	4.9
	135411	L10333	Hs.99947	reticulon 1	5.3
	300023	M10098		AFFX control: 18S ribosomal RNA	4.6
	300254	AW079607	Hs.55610	ESTs; Weakly similar to ZnT-3 [H.sapiens	7.8
	300273	AW013907	Hs.167531	ESTs; Moderately similar to predicted us	11.5

300319	AW157646	Hs.153506	ESTs; Weakly similar to microtubule-acti	8.5
300566	H86709	Hs.326392	son of sevenless (Drosophila) homolog 1	5.8
300578	AI989417	Hs.134289	ESTs	4.4
300671	AI239706	Hs.93810	ESTs	7.9
300675	AA039352	Hs.125034	ESTs; Weakly similar to ORF YDL040c [S.c	4.5
300680	AW468066	Hs.24817	ESTs; Weakly similar to KIAA0986 protein	5.2
300762	AI497778	Hs.20509	ESTs	6.4
300810	AI076890	Hs.146847	ESTs	5.8
300813	AA406411	Hs.208341	ESTs; Weakly similar to KIAA0989 protein	10.6
300823	AI863068	Hs.106823	ESTs; Weakly similar to putative zinc fi	5.6
300834	AF109300	Hs.147924	ESTs	6.7
300923	AW136372	Hs.1852	ESTs	7.6
300962	AA593373	Hs.293744	ESTs	5.5
301015	AA947682	Hs.20252	ESTs; Weakly similar to Chain A; Cdc42hs	7
301042	AI659131	Hs.197733	ESTs	24.9
301242	AW161535	Hs.23782	ESTs	11.8
301254	AI049624	Hs.283390	EST cluster (not in UniGene) with exon h	4.3
301262	H29500	Hs.7130	ESTs; Moderately similar to N-copine [H.	4.3
301388	AA156879	Hs.262036	ESTs; Weakly similar to ZINC FINGER PROT	6.6
301563	AI802946	Hs.44208	ESTs; Weakly similar to match to ESTs AA	5.7
301656	AW008475	Hs.151258	EST cluster (not in UniGene) with exon h	6.8
301689	Z44810	Hs.301789	ESTs; Weakly similar to similar to C.ele	6.3
301783	AL046347	Hs.83937	Homo sapiens PAC clone DJ1159O04 from 7p	6.2
301805	AI800004	Hs.142846	ESTs; Weakly similar to MesP1 [M.musculu	8.5
301846	R20002	Hs.6823	ESTs; Weakly similar to intrinsic factor	4.6
301891	AF131855	Hs.279591	Homo sapiens clone 25056 mRNA sequence	6.3
302005	AI869666	Hs.123119	ESTs	36.8
302056	AI457532	Hs.30488	ESTs; Moderately similar to ROSA26AS [M.	9.5
302067	H05698	Hs.222399	ESTs; Weakly similar to protein-tyrosine	5.8
302099	AL021397	Hs.137576	ribosomal protein L34 pseudogene 1	8.8
302147	AB022660	Hs.151717	KIAA0437 protein	5.9
302214	AJ001454	Hs.159425	Homo sapiens mRNA for testican-3	4.3
302236	AI128606	Hs.6557	zinc finger protein 161	4.3
302358	D81150	Hs.322848	EST cluster (not in UniGene) with exon h	5.5
302410	NM_004917	Hs.218366	EST cluster (not in UniGene) with exon h	26.8
302486	AC003682	Hs.183512	multiple UniGene matches	8.2
302582	NM_000522	Hs.249195	EST cluster (not in UniGene) with exon h	6.4
302785	AA425562	Hs.11065	EST cluster (not in UniGene) with exon h	5
302792	AA343696	Hs.46821	ESTs; Weakly similar to putative [H.sapi	4.8
302881	AA508353	Hs.105314	relaxin 1 (H1)	78.8
302892	N58545	Hs.42346	histone deacetylase 3	8.5
302970	AW118352	Hs.312679	EST cluster (not in UniGene) with exon h	7.4
302977	AW263124	Hs.315111	EST cluster (not in UniGene) with exon h	5.5
303029	AF199613		EST cluster (not in UniGene) with exon h	4.6
303125	AF161352	Hs.111782	EST cluster (not in UniGene) with exon h	5.8
303280	AI571580	Hs.170307	ESTs	4.3
303306	AA215297	Hs.61441	EST cluster (not in UniGene) with exon h	6.4
303309	AL134164	Hs.145416	ESTs	6.6
303344	AA255977	Hs.250646	ESTs; Highly similar to ubiquitin-conjug	19.5
303380	AA298471	Hs.326567	EST cluster (not in UniGene) with exon h	6.6
303401	AA758552	Hs.309497	ESTs	6.8
303525	AW516519	Hs.273294	ESTs	4.8
303526	AA348111	Hs.96900	ESTs	12.1
303540	AA355607	Hs.309490	ESTs; Weakly similar to MMSET type I [H.	8.2
303572	AW338520	Hs.242540	ESTs	8.4
303685	AW500106	Hs.23643	EST cluster (not in UniGene) with exon h	4.9
303699	D30891	Hs.19525	EST cluster (not in UniGene) with exon h	15.7
303702	AW500748	Hs.224961	ESTs; Weakly similar to 73 kDA subunit o	6.3
303718	AI741397	Hs.114658	ESTs	4.6
303722	AA521510	Hs.145010	ESTs	12.5
303732	AW502405	Hs.125759	ESTs; Weakly similar to tumor suppressor	4.3
303735	AA707750	Hs.169055	ESTs; Weakly similar to cis-Golgi matrix	5.4
303752	AI017286	Hs.5957	EST cluster (not in UniGene) with exon h	5.3
303753	AW503733	Hs.9414	ESTs	13
303813	AI275850	Hs.114658	EST cluster (not in UniGene) with exon h	7.8
304053	R00493	Hs.125565	translocase of inner mitochondrial membr	4.8
304218	N66373	Hs.27973	ESTs; Weakly similar to ZK354.7 [C.elega	6
305200	AA668128	Hs.45207	EST singleton (not in UniGene) with exon	5.7
306716	AI024916	Hs.251354	ESTs	5.7

307848	AI364186		EST singleton (not in UniGene) with exon	7.3
307871	AI368665	Hs.31476	EST singleton (not in UniGene) with exon	5.4
308050	AI460004	Hs.31608	EST singleton (not in UniGene) with exon	8.1
308362	AI613519	Hs.105749	EST singleton (not in UniGene) with exon	5.5
308923	AI863051	Hs.279815	ESTs	4.4
309116	AI927149	Hs.29797	ribosomal protein L10	4.5
309375	AW075342	Hs.9271	EST singleton (not in UniGene) with exon	7.4
309674	AW205604	Hs.266009	ESTs; Weakly similar to !!!! ALU SUBFAMI	5
310095	AI921750	Hs.144871	ESTs	5
310098	AI685841	Hs.161354	ESTs	11.6
310250	AI478629	Hs.158465	ESTs	5.8
310365	AI262148	Hs.145569	ESTs	9.7
310382	AI734009	Hs.127699	EST cluster (not in UniGene)	10.4
310409	AI612775	Hs.145710	ESTs	4.6
310431	AI420227	Hs.149358	ESTs	72.9
310573	AW292180	Hs.156142	ESTs	7.6
310598	AI338013	Hs.140546	ESTs	9.2
310639	AW269082	Hs.175162	ESTs	4.5
310787	AW262580	Hs.147674	ESTs	4.9
310816	AI973051	Hs.224965	ESTs	7.6
311251	AI655662	Hs.197698	ESTs	41.3
311280	AI767957	Hs.198248	ESTs; Weakly similar to Y38A8.1 gene pro	4.5
311330	AI679524	Hs.201629	ESTs; Moderately similar to !!!! ALU SUB	4.6
311515	AW136713	Hs.23862	ESTs	5.9
311574	AI824863	Hs.211420	ESTs	4.8
311587	AI828254	Hs.271019	ESTs	5.8
311596	AI682088	Hs.79375	ESTs	26.4
311631	AI809519	Hs.27133	ESTs	6.4
311688	AW025661	Hs.240090	ESTs	7.4
311783	AI682478	Hs.13528	EST	4.6
311826	AA765470	Hs.85092	ESTs	6.7
311853	AW014013	Hs.107056	ESTs	5.3
311901	R16890	Hs.137135	ESTs	5.6
311932	AW451654	Hs.257482	ESTs	4.3
312153	AA759250	Hs.118625	cytochrome b-561	11
312182	AA834800	Hs.326263	EST cluster (not in UniGene)	16.9
312242	AI380207	Hs.125276	ESTs	4.7
312296	C01367	Hs.127128	ESTs	5.3
312407	R46180	Hs.153485	ESTs	6.2
312424	AA847398	Hs.291997	ESTs	4.8
312425	R49353	Hs.293892	ESTs	5.2
312480	R68651	Hs.144997	ESTs	9.5
312518	C17785	Hs.182738	ESTs	6.3
312521	AA033609	Hs.239884	ESTs	11.2
312527	AI695522	Hs.191271	ESTs	4.7
312539	AI004377	Hs.200360	ESTs	7
312546	AI623511	Hs.118567	ESTs	5.1
312563	AA976064	Hs.180842	ESTs	6.5
312623	AA694607	Hs.176956	EST cluster (not in UniGene)	10.8
312857	AA772279	Hs.126914	ESTs	5
312890	AI813654	Hs.5957	ESTs	5.8
312903	AA939266	Hs.278626	ESTs	7.7
312905	H92571	Hs.234478	ESTs	6.5
312976	AA836271	Hs.125830	ESTs	4.6
312983	AI079278	Hs.269899	ESTs	5.1
312996	AA249018	Hs.154331	EST cluster (not in UniGene)	7
313035	N36417	Hs.144928	ESTs	6.3
313166	AI801098	Hs.151500	ESTs	4.3
313188	AI039702	Hs.179573	collagen; type I; alpha 2	4.8
313218	AA827805	Hs.124296	ESTs	5
313226	AI200281	Hs.123910	ESTs	5.9
313325	AI420611	Hs.127832	ESTs	4.6
313326	AI088120	Hs.122329	ESTs	7.4
313425	AA745689	Hs.186838	ESTs; Weakly similar to similar to zinc	6.3
313499	AI261390	Hs.146085	ESTs	5.6
313540	AI797301	Hs.5740	ESTs	5.9
313568	AW467376	Hs.129640	ESTs	4.3
313569	AI273419	Hs.135146	ESTs; Weakly similar to ZK1058.5 [C.eleg	4.6
313603	AW468119	Hs.287631	EST cluster (not in UniGene)	6.8

313615	AW295194	Hs.301997	DKFZP434N126 protein	5.2
313625	AW468402	Hs.254020	ESTs	7.8
313634	AA688292	Hs.337786	ESTs	4.4
313635	AA507227	Hs.6390	ESTs	8.1
313638	AI753075	Hs.104627	ESTs	6.7
313670	C16690	Hs.23767	EST cluster (not in UniGene)	4.4
313671	W49823	Hs.104613	ESTs	4.4
313676	AA861697	Hs.120591	EST cluster (not in UniGene)	13.4
313703	AI161293	Hs.280380	ESTs; Weakly similar to KIAA0525 protein	10
313712	AA768553	Hs.74170	ESTs	5.2
313800	AW296132	Hs.55098	ESTs	5.4
313979	AI535895	Hs.221024	ESTs	4.3
314121	AI732100	Hs.187619	ESTs	13.6
314123	AW245993	Hs.223394	ESTs	6.4
314171	AI821895	Hs.193481	ESTs	29.4
314188	AL138431	Hs.164243	ESTs	4.6
314219	AL036001	Hs.48376	ESTs	5.7
314236	AA743396	Hs.189023	ESTs	4.9
314237	AA732359	Hs.96264	ESTs	4.4
314284	AA731431	Hs.293464	EST cluster (not in UniGene)	6.4
314305	AI280112	Hs.125232	ESTs	5.3
314343	AI754701	Hs.328476	ESTs; Weakly similar to alternatively sp	6.2
314530	AI052358	Hs.193726	ESTs	4.5
314691	AW207206	Hs.136319	ESTs	17
314695	AW502698	Hs.118152	ESTs	8.9
314785	AI538226	Hs.32976	ESTs	9.4
314801	AA481027	Hs.109045	ESTs; Weakly similar to ORF YGR245c [S.c	8
314864	AA493811	Hs.294068	ESTs	6
314907	AI672225	Hs.222886	ESTs	19.3
314916	AA548906	Hs.122244	ESTs	4.5
314954	AA521381	Hs.187726	ESTs	5.3
314981	AA524953	Hs.293334	ESTs	4.6
315021	AA533447	Hs.312989	EST cluster (not in UniGene)	5.1
315051	AW292425	Hs.163484	EST	15.5
315052	AA876910	Hs.134427	ESTs	20
315073	AW452948	Hs.257631	ESTs	5.3
315084	AI821085		ESTs	8.2
315214	AI915927	Hs.34771	ESTs	5.4
315220	AI420753	Hs.66731	ESTs	5.1
315278	AI985544	Hs.12450	ESTs	5.8
315282	AI222165	Hs.144923	ESTs	4.5
315368	AW291563	Hs.104696	ESTs	8
315369	AA764918	Hs.256531	ESTs	4.8
315378	AI263393	Hs.145008	ESTs	6.2
315379	AI378329	Hs.126629	ESTs	5.4
315402	AW293424	Hs.75354	ESTs	5.1
315442	AA977935	Hs.127274	ESTs	6.6
315443	AW003416	Hs.160604	ESTs	5.5
315528	R37257	Hs.184780	ESTs	8.1
315593	AW198103	Hs.158154	ESTs	9.9
315634	AA837085	Hs.220585	ESTs	7.8
315705	AW449285	Hs.313636	ESTs	8.9
315707	AI418055	Hs.161160	ESTs	5.1
315714	AA744015	Hs.298138	EST cluster (not in UniGene)	6.1
315740	T05558	Hs.156880	EST cluster (not in UniGene)	6.8
315762	AI391470	Hs.158618	ESTs	5.3
315769	AA744875	Hs.189413	ESTs	5
315843	AA679430	Hs.191897	ESTs	5.7
315990	AI800041	Hs.190555	ESTs	9.2
316012	AA764950	Hs.119898	ESTs	4.3
316036	AA708016	Hs.190389	ESTs	5.9
316055	AA693880	Hs.6947	EST cluster (not in UniGene)	6.7
316074	AW517542	Hs.293273	ESTs	5.5
316100	AW203986	Hs.213003	ESTs	5.1
316169	AI127483	Hs.120451	ESTs	8.2
316442	AA760894	Hs.153023	ESTs	17.1
316491	AA766025	Hs.186854	EST	4.6
316504	AW135854	Hs.132458	ESTs	4.3
316667	AW015940	Hs.232234	ESTs	7.6

	316854	AA831215	Hs.159066	ESTs; Weakly similar to predicted using	5.1
	316905	AW138241	Hs.210846	ESTs	6.4
	317008	AW051597	Hs.143707	ESTs	4.4
	317019	AA864968	Hs.127699	ESTs	11
5	317194	AW445167	Hs.126036	ESTs	13.5
	317224	D56760	Hs.93029	ESTs	8.7
	317404	AI806867	Hs.126594	ESTs	8.7
	317501	AA931245	Hs.137097	ESTs	11.1
	317548	AI654187	Hs.195704	ESTs	14.2
10	317651	AW292779	Hs.169799	ESTs	5.8
	317758	AI733277	Hs.128321	ESTs	5.4
	317850	N29974	Hs.152982	EST cluster (not in UniGene)	11.4
	317869	AW295184	Hs.129142	ESTs; Weakly similar to DEOXYRIBONUCLEAS	13.8
	317902	AI828602	Hs.211265	ESTs	5.3
15	317916	AI565071	Hs.159983	ESTs	7.7
	318239	AI085198	Hs.164226	ESTs	13.1
	318268	AI817736	Hs.182490	ESTs	6.2
	318327	AW294013	Hs.200942	ESTs	4.6
	318363	R45530	Hs.1440	gamma-aminobutyric acid (GABA) A recepto	6
20	318428	AI949409	Hs.194591	ESTs	12.3
	318464	AI151010	Hs.157774	ESTs	4.3
	318524	AW291511	Hs.159066	ESTs	25.9
	318540	T30280	Hs.274803	EST cluster (not in UniGene)	7
	318591	AW206806	Hs.115325	ESTs	4.8
25	318615	AI133617	Hs.10177	ESTs	5.5
	318646	AW175665	Hs.278695	ESTs	5.7
	318667	AI493742	Hs.165210	ESTs	11
	318668	W26276	Hs.136075	ESTs	5.9
	318753	AA578265	Hs.7130	copine IV	5.5
30	319080	Z45131	Hs.23023	ESTs	16.9
	319181	F06504	Hs.27384	EST cluster (not in UniGene)	4.6
	319191	AF071538	Hs.79414	prostate epithelium-specific Ets transcr	6.6
	319233	R21054	Hs.180532	ESTs	4.9
	319586	D78608	Hs.283683	ESTs	8.2
35	319750	AA621606	Hs.117956	ESTs	9.3
	319763	AA460775	Hs.6295	ESTs	14.3
	319824	AA424266	Hs.123642	EST cluster (not in UniGene)	12.8
	319838	AA337642	Hs.95262	nuclear factor related to kappa B bindin	5.1
	319913	AA179304	Hs.271586	ESTs; Moderately similar to !!!! ALU SUB	4.3
40	319964	T80579	Hs.290270	ESTs	5.8
	320076	AI653733	Hs.271593	ESTs	8.5
	320102	AW296219	Hs.115325	RAB7; member RAS oncogene family-like 1	9.8
	320187	T99949	Hs.303428	EST cluster (not in UniGene)	9.8
	320211	AL039402	Hs.125783	DEME-6 protein	7.9
45	320324	AF071202	Hs.139336	ATP-binding cassette; sub-family C (CFTR	56.2
	320455	R49889	Hs.24144	EST cluster (not in UniGene)	8.3
	320464	AI089817	Hs.237146	ESTs	5.4
	320561	NM_006953	Hs.159330	EST cluster (not in UniGene)	7
	320574	AL049443	Hs.161283	Homo sapiens mRNA; cDNA DKFZp586N2020 (f	4.4
50	320576	AL049977	Hs.162209	Homo sapiens mRNA; cDNA DKFZp564C122 (fr	6.7
	320654	AW263086	Hs.118112	ESTs	6
	320796	AF038966	Hs.31218	secretory carrier membrane protein 1	13.5
	320800	AI681006	Hs.71721	ESTs	6.2
	320813	AW360847	Hs.16578	ESTs	9.3
55	320853	AI473796	Hs.135904	ESTs	8.1
	320856	D59945	Hs.65366	EST cluster (not in UniGene)	6
	320899	AA633772	Hs.116796	ESTs	9.2
	320918	AW195012	Hs.293970	ESTs	5
	320973	H19732	Hs.247917	ESTs	5.9
60	321099	AA018386	Hs.64341	ESTs	4.6
	321190	H52462	Hs.163872	EST cluster (not in UniGene)	5.8
	321318	AB033041	Hs.137507	EST cluster (not in UniGene)	8.4
	321382	AW372449	Hs.175982	EST cluster (not in UniGene)	7.3
	321441	AW297633	Hs.118498	ESTs	14.7
65	321538	H80483	Hs.46903	EST cluster (not in UniGene)	9.2
	321609	H86021	Hs.182538	ESTs; Weakly similar to hMmTRA1b [H.sapi	4.8
	321636	AI791838	Hs.193465	ESTs	5.5
	321638	AI356352	Hs.108932	ESTs	4.6
	321644	AI204177	Hs.237396	ESTs	6.6

321681	AA233821	Hs.190173	EST cluster (not in UniGene)	4.6
321726	X91221	Hs.144465	EST cluster (not in UniGene)	5
321758	U29112	Hs.196151	EST cluster (not in UniGene)	6.2
321877	AL109784	Hs.189222	EST cluster (not in UniGene)	4.6
321899	N55158	Hs.29468	ESTs	4.6
321902	AA746374	Hs.145010	ESTs	8.2
322007	AW410646	Hs.164649	ESTs	5.1
322055	AL137646	Hs.146001	EST cluster (not in UniGene)	4.3
322092	AF085833	Hs.135624	EST cluster (not in UniGene)	4.3
322221	AI690619	Hs.179662	nucleosome assembly protein 1-like 1	4.4
322278	AF086283		EST cluster (not in UniGene)	5.8
322303	W07459	Hs.157601	EST cluster (not in UniGene)	22
322437	AW393804	Hs.170253	ESTs; Weakly similar to rabaptin-4 [H.sa	4.4
322493	AF143235	Hs.279819	EST cluster (not in UniGene)	7.2
322782	AA056060	Hs.202577	EST cluster (not in UniGene)	18.4
322811	AA782292	Hs.105872	ESTs	6.9
322818	AW043782	Hs.293616	ESTs	10.7
322826	AI807883	Hs.180059	ESTs	5
322887	AI986306	Hs.86149	ESTs; Weakly similar to KIAA0969 protein	11.9
322889	AA081924	Hs.124918	ESTs	7.1
322924	AA669253	Hs.136075	ESTs	4.5
322982	AI351191	Hs.128430	ESTs	6.6
322994	AA422116	Hs.191461	ESTs	4.7
323040	AA336609	Hs.10862	ESTs	6.9
323041	AL118747	Hs.26691	EST cluster (not in UniGene)	8.3
323045	AA148950	Hs.188836	ESTs	4.6
323048	AL118923	Hs.175110	EST cluster (not in UniGene)	7.5
323070	AA157726	Hs.264330	ESTs	7.5
323071	AA157867	Hs.5722	ESTs	4.7
323097	Z44354	Hs.296261	guanine nucleotide binding protein (G pr	4.9
323131	AA176982	Hs.270124	EST cluster (not in UniGene)	6.1
323136	AL120351	Hs.30177	EST cluster (not in UniGene)	4.3
323175	AI827137	Hs.336454	ESTs	6.2
323218	AF131846	Hs.13396	Homo sapiens clone 25028 mRNA sequence	6.3
323226	AF055019	Hs.21906	Homo sapiens clone 24670 mRNA sequence	12.6
323236	AA363148	Hs.293960	ESTs	10.9
323262	AI829770	Hs.190642	ESTs	7.6
323276	AA836452	Hs.323822	ESTs	7.6
323287	AA639902	Hs.104215	ESTs	24.7
323335	AI655499	Hs.161712	ESTs	14.1
323341	AL134875	Hs.108646	ESTs	5.3
323362	AL135067	Hs.117182	ESTs	6.1
323486	C05278	Hs.299221	ESTs; Moderately similar to [PYRUVATE DE	8.5
323496	AI826801	Hs.300700	ESTs	4.5
323507	H71721	Hs.128387	ESTs	4.4
323545	AI814405	Hs.224569	ESTs	5.8
323623	AA314280	Hs.146589	EST cluster (not in UniGene)	5
323663	AW263526	Hs.243023	ESTs	7.7
323691	AA317561	Hs.145599	EST cluster (not in UniGene)	5.9
323810	AA740405	Hs.108806	ESTs	6.2
323846	AA337621	Hs.137635	ESTs	6
323929	AA354940	Hs.145958	ESTs	10.7
323959	AI636775	Hs.6831	ESTs	5.4
323996	AA367032	Hs.217882	ESTs	5.8
323997	AA844907	Hs.274454	EST cluster (not in UniGene)	4.4
324019	AW177009		EST cluster (not in UniGene)	4.6
324130	AL046575	Hs.130198	ESTs	11
324295	AI146686	Hs.143691	ESTs	13.7
324296	AI524039	Hs.192524	ESTs	6.8
324307	AA627642	Hs.4994	transducer of ERBB2; 2 (TOB2)	4.9
324330	AA884766		EST cluster (not in UniGene)	4.3
324385	F28212	Hs.284247	EST cluster (not in UniGene)	4.7
324430	AA464018	Hs.184598	EST cluster (not in UniGene)	13.6
324452	AW014022	Hs.170953	ESTs	7.6
324547	AW501974	Hs.74170	ESTs	5.6
324603	AW016378	Hs.292934	ESTs	24.2
324617	AA508552	Hs.195839	ESTs	54
324618	AI346282	Hs.87159	ESTs	4.6
324620	AA448021	Hs.94109	EST cluster (not in UniGene)	5.7

324626	AI685464		ESTs	9	
324658	AI694767	Hs.129179	ESTs	22	
324676	AW503943	Hs.112451	ESTs	4.9	
324691	AI217963	Hs.293341	ESTs; Weakly similar to Pro-a2(XI) [H.sa	10.6	
324696	AA641092	Hs.257339	ESTs	10.2	
324713	AW340249	Hs.163440	ESTs	5.5	
324715	AI739168	Hs.131798	EST cluster (not in UniGene)	7.2	
324718	AI557019	Hs.116467	ESTs	34.4	
324720	AA578904	Hs.292437	ESTs	4.8	
324752	AI279919	Hs.272072	ESTs; Moderately similar to !!!! ALU SUB	7.9	
324753	AA612626	Hs.144871	EST cluster (not in UniGene)	5.2	
324790	AI334367	Hs.159337	ESTs	7.6	
324801	AI819924	Hs.14553	ESTs	12.6	
324804	AI692552		ESTs	6.5	
324845	AA361016	Hs.337533	ESTs	4.5	
324888	AI564134	Hs.136102	KIAA0853 protein	4.4	
324929	AI741633	Hs.125350	ESTs	6.5	
324961	AA613792		EST cluster (not in UniGene)	5.1	
325108	AA401863	Hs.22380	ESTs	7.1	
326816			CH.20_hs gi|6552458	9.6	
326997			CH.21_hs gi|5867660	4.8	
327098			CH.21_hs gi|6682516	4.3	
328492			CH.07_hs gi|5868455	5.8	
329362			CH.X_hs gi|5868837	4.3	
329929			CH.16_p2 gi|6165201	5.5	
329960			CH.16_p2 gi|5091594	7.6	
330020			CH.16_p2 gi|6671887	6	
330211			CH.05_p2 gi|6013592	12.6	
330384	M23263		androgen receptor (dihydrotestosterone r	9	
330430	HG2261-HT2352		Hs.321110	Antigen, Prostate Specific, Alt. Splice	13.8
330546	U31382	Hs.299867	guanine nucleotide binding protein 4	6	
330551	U39840		hepatocyte nuclear factor 3; alpha	4.9	
330658	AA319514	Hs.30732	ESTs	6	
330700	AA037415	Hs.20999	ESTs	5.5	
330704	AA056557	Hs.6759	ESTs	5.1	
330705	AA102571	Hs.157078	ESTs	11.7	
330706	AA121140	Hs.177576	ESTs; Moderately similar to kynurenine a	14.5	
330712	AA167269	Hs.52620	ESTs	5	
330725	AA252033	Hs.24052	ESTs; Weakly similar to !!!! ALU SUBFAMI	7.2	
330732	AA281092	Hs.35254	ESTs	4.9	
330762	AA449677	Hs.15251	Human DNA sequence from clone 437M21 on	18.5	
330763	AA450200	Hs.143187	FK506-binding protein 3 (25kD)	4.3	
330772	AA479114	Hs.11356	ESTs	5.8	
330786	D60374		EST	4.6	
330892	AA149579	Hs.91202	ESTs	15.3	
330949	H01458	Hs.142896	ESTs	10.3	
330977	H20826	Hs.315181	ESTs	4.4	
331017	N24619	Hs.108920	ESTs	11.8	
331099	R36671	Hs.14846	ESTs	11.6	
331128	R51361	Hs.268714	ESTs	4.8	
331151	R82331	Hs.268838	ESTs	13	
331195	T64447	Hs.168439	ESTs	4.9	
331320	AA262999	Hs.300141	ESTs	4.8	
331321	AA278355	Hs.87929	ESTs	6.1	
331337	AA287662	Hs.118630	ESTs	9.2	
331348	AA400596	Hs.88143	ESTs	9.9	
331359	AA416979	Hs.81897	ESTs	4.3	
331383	AA454543	Hs.43543	ESTs	4.6	
331422	F10802	Hs.237339	ESTs; Moderately similar to !!!! ALU SUB	4.9	
331442	H77381	Hs.41223	ESTs	7.5	
331466	N21680	Hs.43455	ESTs	5.4	
331479	N27154	Hs.44076	ESTs	6.5	
331490	N32912	Hs.291039	ESTs; Weakly similar to hypothetical 43.	12.5	
331493	N34357	Hs.93817	ESTs	4.6	
331561	N62780	Hs.48703	ESTs	9.2	
331615	N92352	Hs.5472	ESTs	4.6	
331659	W48868	Hs.334305	ESTs	8.7	
331696	Z38907	Hs.65949	KIAA0888 protein	10.3	
331811	AA404500	Hs.187958	ESTs	4.8	

331848	AA417039	Hs.98268	signal recognition particle 72kD	7.5
331873	AA429445	Hs.98640	ESTs	6.5
331889	AA431407	Hs.98802	Homo sapiens Chromosome 16 BAC clone CIT	33.6
331967	AA460158	Hs.99589	KIAA1028 protein	6.8
331974	AA464518	Hs.105322	ESTs	5.3
332043	AA490831	Hs.201591	ESTs	10.8
332076	AA599477	Hs.291156	ESTs	4.4
332173	F09281	Hs.100725	ESTs	5.5
332247	N58172		ESTs	14.2
332249	N62096	Hs.194140	ESTs	7.2
332325	T79428	Hs.339667	ESTs	5.6
332396	AA340504		ESTs; Weakly similar to similarto human	21.2
332434	N75542	Hs.237731	transcription factor 4	15.3
332493	N95495	Hs.56729	ESTs; Highly similar to GTP-binding prot	7.1
332522	L38503	Hs.178357	glutathione S-transferase theta 2	6.6
332526	AA281753	Hs.17731	inositol 1;4;5-triphosphate receptor; ty	5.8
332530	M31682	Hs.19280	inhibin; beta B (activin AB beta polypep	5.5
332533	M99487	Hs.325825	folate hydrolase (prostate-specific memb	38.1
332538	N48715	Hs.20991	ESTs	6.5
332546	D84454	Hs.22587	solute carrier family 35 (UDP-galactose	4.8
332594	AA279313	Hs.32951	methyl CpG binding protein 2	5.6
332610	AA412405	Hs.40513	ESTs; Weakly similar to BETA GALACTOSIDA	5.6
332661	N95742	Hs.6390	ESTs	6.9
332697	T94885	Hs.75725	carboxypeptidase E	24.3
332712	D26070	Hs.79306	inositol 1;4;5-triphosphate receptor; ty	9.9
332716	L00058	Hs.79630	v-myc avian myelocytomatosis viral oncog	5.6
332726	R72029	Hs.83428	synaptophysin-like protein	5
332781	AA233258		ESTs; Weakly similar to D1007.5 [C.elega	4.5
332797			CH22_FGENES.6_2	30.8
332798			CH22_FGENES.6_5	66.8
332799			CH22_FGENES.6_6	19.8
332933			CH22_FGENES.38_7	5.6
332980			CH22_FGENES.54_1	5.5
332984			CH22_FGENES.54_6	4.9
333168			CH22_FGENES.94_1	4.7
333169			CH22_FGENES.94_2	4.4
333452			CH22_FGENES.157_1	4.8
333456			CH22_FGENES.157_5	4.3
333458			CH22_FGENES.157_7	4.6
333611			CH22_FGENES.217_6	4.7
333621			CH22_FGENES.219_5	5.5
333814			CH22_FGENES.282_2	7.1
333849			CH22_FGENES.290_8	6.2
333949			CH22_FGENES.303_5	4.3
333951			CH22_FGENES.303_7	4.9
333955			CH22_FGENES.303_11	5.6
334150			CH22_FGENES.339_1	5.1
334223			CH22_FGENES.360_4	20.3
334297			CH22_FGENES.372_3	9.4
334443			CH22_FGENES.387_2	4.6
334444			CH22_FGENES.387_4	5.6
334447			CH22_FGENES.387_7	13.1
334570			CH22_FGENES.405_11	5.4
334749			CH22_FGENES.427_1	5.3
334777			CH22_FGENES.430_9	4.7
334960			CH22_FGENES.465_29	5.2
335179			CH22_FGENES.504_9	8.8
335293			CH22_FGENES.527_6	4.7
335550			CH22_FGENES.576_11	5.1
335581			CH22_FGENES.581_19	5.7
335586			CH22_FGENES.581_25	4.3
335809			CH22_FGENES.617_6	6.2
335810			CH22_FGENES.617_7	5.8
335822			CH22_FGENES.619_7	7.1
335824			CH22_FGENES.619_11	8.5
335853			CH22_FGENES.626_5	4.3
335886			CH22_FGENES.632_4	4.3
336034			CH22_FGENES.678_5	6.8
336441			CH22_FGENES.827_7	7.6

336624	CH22_FGENES.6-3	43.3
336625	CH22_FGENES.6-4	37.9
336679	CH22_FGENES.43-7	5.3
337577	CH22_C65E1.GENSCAN.8-1	4.9
338255	CH22_EM:AC005500.GENSCAN.276-3	13.4
338260	CH22_EM:AC005500.GENSCAN.279-10	4.6
338561	CH22_EM:AC005500.GENSCAN.421-5	4.6
338562	CH22_EM:AC005500.GENSCAN.421-6	4.3
338759	CH22_EM:AC005500.GENSCAN.517-6	5.1
338763	CH22_EM:AC005500.GENSCAN.517-16	5.5
338764	CH22_EM:AC005500.GENSCAN.517-17	7.1

TABLE 3A shows the accession numbers for those primekeys lacking unigeneID's for Table 3. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Pkey: Unique Eos probeset identifier number
CAT number: Gene cluster number
Accession: Genbank accession numbers

Pkey	CAT number	Accession
123619	371681_1	AA602964 AA609200
116722	143512_1	Z24878 AA494098 F13654 AA494040 AA143127
103677	41847_1	Z83806 AJ132091 AJ132090
125992	1589048_1	H48372 W01626
109342	genbank_AA213620	AA213620
125154	genbank_W38419	W38419
101447	entrez_M21305	M21305
124357	genbank_N22401	N22401
108910	genbank_AA136590	AA136590
322278	47271_1	W69304 AF086283 W69200
315084	350959_1	AI821085 AW973464 AA554802 AI821831 AA657438 AA640756 AA650339
324019	262792_1	AW177009 AI381610
324330	300543_1	AA884766 AW974271 AA592975 AA447312
324626	336411_1	AI685464 AW971336 AA513587 AA525142
303029	37699_1	AF199613 AF108756
324804	398093_1	AI692552 AI393343 AI800510 AI377711 F24263 AA661876
324961	376239_1	AA613792 AW182329 T05304 AW858385
329362	c_x_hs	
336624	CH22_4071FG_6_3_	
336625	CH22_4072FG_6_4_	
336679	CH22_4157FG_43_7_	
338255	CH22_6856FG__LINK_EM:AC00	
338260	CH22_6863FG__LINK_EM:AC00	
329929	c16_p2	
329960	c16_p2	
338561	CH22_7294FG__LINK_EM:AC00	
338562	CH22_7295FG__LINK_EM:AC00	
338759	CH22_7581FG__LINK_EM:AC00	
338763	CH22_7585FG__LINK_EM:AC00	
338764	CH22_7586FG__LINK_EM:AC00	
333168	CH22_400FG_94_1_LINK_EM:A	
333169	CH22_401FG_94_2_LINK_EM:A	
333452	CH22_702FG_157_1_LINK_EM:	
333456	CH22_706FG_157_5_LINK_EM:	
333458	CH22_708FG_157_7_LINK_EM:	
333611	CH22_872FG_217_6_LINK_EM:	
333621	CH22_882FG_219_5_LINK_EM:	
333814	CH22_1083FG_282_2_LINK_EM	
333849	CH22_1118FG_290_8_LINK_EM	
335179	CH22_2515FG_504_9_LINK_EM	
333949	CH22_1225FG_303_5_LINK_EM	
333951	CH22_1227FG_303_7_LINK_EM	
333955	CH22_1231FG_303_11_LINK_E	
335293	CH22_2635FG_527_6_LINK_EM	
326816	c20_hs	
326997	c21_hs	
335550	CH22_2905FG_576_11_LINK_E	
335581	CH22_2938FG_581_19_LINK_E	
335586	CH22_2944FG_581_25_LINK_E	

328492	c_7_hs	
335809	CH22_3181FG_617_6_LINK_EM	
335810	CH22_3182FG_617_7_LINK_EM	
335822	CH22_3195FG_619_7_LINK_EM	
335824	CH22_3197FG_619_11_LINK_E	
335853	CH22_3228FG_626_5_LINK_EM	
335886	CH22_3261FG_632_4_LINK_EM	
330020	c16_p2	
330211	c_5_p2	
337577	CH22_5864FG__LINK_C65E1.G	
307848	AI364186	
332797	CH22_13FG_6_2_LINK_C4G1.G	
332798	CH22_14FG_6_5_LINK_C4G1.G	
332799	CH22_15FG_6_6_LINK_C4G1.G	
334150	CH22_1429FG_339_1_LINK_EM	
332933	CH22_154FG_38_7_LINK_C20H	
332980	CH22_204FG_54_1_LINK_EM:A	
332984	CH22_208FG_54_6_LINK_EM:A	
334223	CH22_1507FG_360_4_LINK_EM	
334297	CH22_1588FG_372_3_LINK_EM	
327098	c21_hs	
334443	CH22_1742FG_387_2_LINK_EM	
334444	CH22_1743FG_387_4_LINK_EM	
334447	CH22_1746FG_387_7_LINK_EM	
334570	CH22_1875FG_405_11_LINK_E	
334749	CH22_2061FG_427_1_LINK_EM	
334777	CH22_2089FG_430_9_LINK_EM	
336034	CH22_3419FG_678_5_LINK_DJ	
334960	CH22_2281FG_465_29_LINK_E	
336441	CH22_3861FG_827_7_LINK_DJ	
330551	9851_2	U39840 NM_004496 AW135607 BE087458 BE087567 AA177116 AW195705 AW750756 AI811008 AI694151 BE348594 AW971075 AI347950 AI201455 AI073898 AA652680 AA613671 AI318364 AA507550 AA693692 AI032599 AA991871 AI269801 AW948974 T74639 AA532907 AW949173
330786	53973_3	BE379594 AI192455 AL039862 AI744012 AI761735 AW243181 AI743687 AI928223 AI423022 AI627855 AI636059 AI651571 AW802044 AI826995 AI431733 AI539125 AA863056 AW270910 AI768930 AW008835 AW615183 AW591147 AI695294 AI672106 AA506358 AI308060 AA011556 AA962437 AI935488 BE219625 AI004356 AW151394 AI218466 N66178 AI419784 AW242519 AW946907 D60374 AA989263 AI698799 AA470460 AI824167
332247	372959_1	AA669097 AA513815 AA026798 AA676526 AA704429 AA704269 AW118292 AA579216 N59172
332396	20265_1	AW579842 BE156562 BE156690 BE156489 BE081033 AK001559 BE149402 M85387 AW367811 AW367798 R17370 AI908947 AA382932 R58449 H18732 AA371231 AW962899 AA713530 AW892946 R53463 H11063 AW068542 Z40761 BE176212 BE176155 W23952 W92188 AW374883 AA303497 AW954769 AA036808 BE168063 AW382073 AW382085 AL041475 H80748 AI078161 BE463983 AI805213 AI761264 W94885 N94502 AI623772 AI419532 AI810302 AI634190 AW002516 AW150777 AI352312 AI367474 AW204807 AI675502 AI337026 AW134715 BE328451 AI123157 AI560020 AI300745 AI608631 AI248873 AA742484 AW051635 H18646 AI245045 AA507111 AI640510 AI925594 AA115747 AA143035 AA151106
332781	32044_1	AK001764 BE313896 AA380199 AA380151 AA194996 AW118089 AA495871 AW975219 AW085598 AI378909 AW992310 AW992409 AI911857 AA657643 AI804471 AI242589 AI623968 R09556 AI129100 AI206500 AA680094 AA677784 AI023178 AI277519 AA424742 AI240654 AA232846 AI804273 AI382376 AA001729 W90790 BE090656 AW295015 AI674596 AI431734 AI420517 AW769185 AI128355 AI192474 AI820001 AA001929 AA706925 AI076676 AI499119 AI200493 AI695919 AI376217 W69195 W69261 AW305099 W90320 BE048357 AI658856 AA838534 AA233258 AI753393 AA709227 AI674387 AI872616

TABLE 3B shows the genomic positioning for those primekeys lacking unigene ID's and accession numbers in Table 3. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

Pkey: Unique number corresponding to an Eos probeset
Ref: Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-495.
Strand: Indicates DNA strand from which exons were predicted.
Nt_position: Indicates nucleotide positions of predicted exons.

Pkey	Ref	Strand	Nt_position
333611	Dunham, I. et.al.	Plus	6548368-6548507
333621	Dunham, I. et.al.	Plus	8597414-8597560
333814	Dunham, I. et.al.	Plus	7894165-7894252
333849	Dunham, I. et.al.	Plus	8018323-8018472
333949	Dunham, I. et.al.	Plus	8589634-8589791
333951	Dunham, I. et.al.	Plus	8592501-8592637
333955	Dunham, I. et.al.	Plus	8597414-8597560
334150	Dunham, I. et.al.	Plus	10529221-10529854
334297	Dunham, I. et.al.	Plus	13420934-13421058
334443	Dunham, I. et.al.	Plus	14298981-14299056
334444	Dunham, I. et.al.	Plus	14306433-14306492
334447	Dunham, I. et.al.	Plus	14308764-14308824
334570	Dunham, I. et.al.	Plus	14994868-14994943
334777	Dunham, I. et.al.	Plus	16259586-16260166
335179	Dunham, I. et.al.	Plus	21634405-21634526
335581	Dunham, I. et.al.	Plus	24976198-24976334
335586	Dunham, I. et.al.	Plus	24990333-24990497
335809	Dunham, I. et.al.	Plus	26310772-26310909
335810	Dunham, I. et.al.	Plus	26314767-26314849
335822	Dunham, I. et.al.	Plus	26364087-26364196
335824	Dunham, I. et.al.	Plus	26376860-26376942
335886	Dunham, I. et.al.	Plus	26934235-26934364
336034	Dunham, I. et.al.	Plus	29014404-29014590
336441	Dunham, I. et.al.	Plus	34187606-34187663
337577	Dunham, I. et.al.	Plus	595377-595678
338260	Dunham, I. et.al.	Plus	15458919-15459257
332797	Dunham, I. et.al.	Minus	216964-216798
332798	Dunham, I. et.al.	Minus	232147-231974
332799	Dunham, I. et.al.	Minus	232421-232307
332933	Dunham, I. et.al.	Minus	2035790-2035681
332980	Dunham, I. et.al.	Minus	5136165-5136019
332984	Dunham, I. et.al.	Minus	2632606-2632457
333168	Dunham, I. et.al.	Minus	3729896-3729788
333169	Dunham, I. et.al.	Minus	3730864-3730767
333452	Dunham, I. et.al.	Minus	5136165-5136019
333456	Dunham, I. et.al.	Minus	2631933-2631797
333458	Dunham, I. et.al.	Minus	5143942-5143806
334223	Dunham, I. et.al.	Minus	12734365-12734269
334749	Dunham, I. et.al.	Minus	16090686-16090106
334960	Dunham, I. et.al.	Minus	20160968-20160795
335293	Dunham, I. et.al.	Minus	22316408-22316275
335550	Dunham, I. et.al.	Minus	24668714-24668658
335853	Dunham, I. et.al.	Minus	26614629-26614506
336624	Dunham, I. et.al.	Minus	227714-227577
336625	Dunham, I. et.al.	Minus	229124-229024
336679	Dunham, I. et.al.	Minus	2035790-2035681
338255	Dunham, I. et.al.	Minus	15242294-15242231
338561	Dunham, I. et.al.	Minus	22311966-22311856
338562	Dunham, I. et.al.	Minus	22312594-22312465
338759	Dunham, I. et.al.	Minus	26582475-26582199
338763	Dunham, I. et.al.	Minus	26628148-26628009
338764	Dunham, I. et.al.	Minus	26641232-26641101

329960	5091594	Minus	1031-1162
329929	6165201	Minus	156410-156553
330020	6671887	Plus	172397-172491
326816	6552458	Plus	198354-198436
326997	5867660	Minus	71389-72147
327098	6682516	Minus	1061684-1062361
330211	6013592	Plus	59158-59215
328492	5868455	Minus	46094-46241
329362	5868837	Minus	65688-68173

TABLE 4: shows a preferred subset of the Accession numbers for genes found in Table 3 which are differentially expressed in prostate tumor tissue compared to normal prostate tissue.

Pkey:	Unique Eos probeset identifier number
ExAccn:	Exemplar Accession number, Genbank accession number
UnigeneID:	Unigene number
Unigene Title:	Unigene gene title
R1:	Ratio of tumor to normal body tissue

Pkey	ExAccn	UnigeneID	Unigene Title	R1
100819	HG4020-HT4290	Hs.2387	Transglutaminase	10.5
102698	U75272	Hs.1867	progastricsin (pepsinogen C)	10.6
102869	X02544	Hs.572	orosomucoid 1	22.6
105370	AA236476	Hs.22791	ESTs; Weakly similar to transmembrane pr	10.3
105645	AA282138	Hs.11325	ESTs	14
106094	AA419461	Hs.23317	ESTs	10.9
109014	AA156790	Hs.262036	ESTs	15.3
109562	F01811	Hs.187931	ESTs; Moderately similar to voltage-gate	10.8
113021	T23855	Hs.129836	KIAA1028 protein	10.8
114124	Z38595	Hs.125019	ESTs; Highly similar to KIAA0886 protein	21.3
122791	AA460158	Hs.129836	KIAA1028 protein	12.4
124352	N21626	Hs.102406	ESTs	10.2
301042	AI659131	Hs.197733	ESTs	24.9
302005	AI869666	Hs.123119	ESTs	36.8
302410	NM_004917	Hs.218366	EST cluster (not in UniGene) with exon h	26.8
302881	AA508353	Hs.105314	relaxin 1 (H1)	78.8
303344	AA255977	Hs.250646	ESTs; Highly similar to ubiquitin-conjug	19.5
303753	AW503733	Hs.9414	ESTs	13
310431	AI420227	Hs.149358	ESTs	72.9
311251	AI655662	Hs.197698	ESTs	41.3
311596	AI682088	Hs.79375	ESTs	26.4
312153	AA759250	Hs.118625	cytochrome b-561	11
312521	AA033609	Hs.239884	ESTs	11.2
313676	AA861697	Hs.120591	EST cluster (not in UniGene)	13.4
314171	AI821895	Hs.193481	ESTs	29.4
314907	AI672225	Hs.222686	ESTs	19.3
315051	AW292425	Hs.163484	EST	15.5
315052	AA876910	Hs.134427	ESTs	20
317548	AI654187	Hs.195704	ESTs	14.2
317869	AW295184	Hs.129142	ESTs; Weakly similar to DEOXYRIBONUCLEAS	13.8
318428	AI949409	Hs.194591	ESTs	12.3
318524	AW291511	Hs.159066	ESTs	25.9
319080	Z45131	Hs.23023	ESTs	16.9
319763	AA460775	Hs.6295	ESTs	14.3
320324	AF071202	Hs.139336	ATP-binding cassette; sub-family C (CFTR	56.2
321441	AW297633	Hs.118498	ESTs	14.7
322303	W07459	Hs.157601	EST cluster (not in UniGene)	22
322782	AA056060	Hs.202577	EST cluster (not in UniGene)	18.4
322818	AW043782	Hs.293616	ESTs	10.7
323287	AA639902	Hs.104215	ESTs	24.7
324603	AW016378	Hs.292934	ESTs	24.2
324617	AA508552	Hs.195839	ESTs	54
324658	AI694767	Hs.129179	ESTs	22
324691	AI217963	Hs.293341	ESTs; Weakly similar to Pro-a2(XI) [H.sa	10.6
324696	AA641092	Hs.257339	ESTs	10.2
324718	AI557019	Hs.116467	ESTs	34.4
330211			CH.05_p2 gi|6013592	12.6
330430	HG2261-HT2352	Hs.321110	Antigen, Prostate Specific, Alt. Splice	13.8
330706	AA121140	Hs.177576	ESTs; Moderately similar to kynurenine a	14.5
330762	AA449677	Hs.15251	Human DNA sequence from clone 437M21 on	18.5
330892	AA149579	Hs.91202	ESTs	15.3
330949	H01458	Hs.142896	ESTs	10.3

	331099	R36671	Hs.14846	ESTs	11.6
	331151	R82331	Hs.268838	ESTs	13
	331889	AA431407	Hs.98802	Homo sapiens Chromosome 16 BAC clone CIT	33.6
	332247	N58172		ESTs	14.2
5	332396	AA340504		ESTs; Weakly similar to similarto human	21.2
	332533	M99487	Hs.325825	folate hydrolase (prostate-specific memb	38.1
	332697	T94885	Hs.75725	carboxypeptidase E	24.3
	332797			CH22_FGENES.6_2	30.8
	332798			CH22_FGENES.6_5	66.8
10	332799			CH22_FGENES.6_6	19.8
	334223			CH22_FGENES.360_4	20.3
	336624			CH22_FGENES.6-3	43.3
	336625			CH22_FGENES.6-4	37.9

TABLE 4A shows the accession numbers for those primekeys lacking unigeneID's for Table 4. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Pkey: Unique Eos probeset identifier number
CAT number: Gene cluster number
Accession: Genbank accession numbers

Pkey	CAT number	Accession
336624	CH22_4071FG_6_3_	
336625	CH22_4072FG_6_4_	
330211	c_5_p2	
332797	CH22_13FG_6_2_LINK_C4G1.G	
332798	CH22_14FG_6_5_LINK_C4G1.G	
332799	CH22_15FG_6_6_LINK_C4G1.G	
334223	CH22_1507FG_360_4_LINK_EM	
332247	372969_1	AA669097 AA513815 AA026798 AA676526 AA704429 AA704269 AW118292 AA579216 N58172
332396	20265_1	AW579842 BE156562 BE156690 BE156489 BE081033 AK001559 BE149402 M85387 AW367811 AW367798 R17370 AI908947 AA362932 R58449 H18732 AA371231 AW962899 AA713530 AW892946 R53483 H11063 AW068542 Z40761 BE176212 BE176155 W23952 W92188 AW374883 AA303497 AW954769 AA036808 BE168063 AW382073 AW382085 AL041475 H80748 AI078161 BE463983 AI805213 AI761264 W94885 N94502 AI623772 AI419532 AI810302 AI634190 AW002516 AW150777 AI352312 AI367474 AW204807 AI675502 AI337026 AW134715 BE328451 AI123157 AI560020 AI300745 AI608631 AI248873 AA742484 AW051635 H18646 AI245045 AA507111 AI640510 AI925594 AA115747 AA143035 AA151106

TABLE 4B shows the genomic positioning for those primekeys lacking unigene ID's and accession numbers in Table 4. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

Pkey:	Unique number corresponding to an Eos probeset
Ref:	Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-495.
Strand:	Indicates DNA strand from which exons were predicted.
Nt_position:	Indicates nucleotide positions of predicted exons.

Pkey	Ref	Strand	Nt_position
332797	Dunham, I. et.al.	Minus	216964-216798
332798	Dunham, I. et.al.	Minus	232147-231974
332799	Dunham, I. et.al.	Minus	232421-232307
334223	Dunham, I. et.al.	Minus	12734365-12734269
336624	Dunham, I. et.al.	Minus	227714-227577
336625	Dunham, I. et.al.	Minus	229124-229024
330211	6013592	Plus	59158-59215

TABLE 5: 1170 GENES UP-REGULATED IN PROSTATE CANCER COMPARED TO NORMAL ADULT TISSUES

Table 5 shows 1170 genes up-regulated in prostate cancer compared to normal adult tissues. These were selected from 59680 probesets on the Affymetrix/Eos Hu03 GeneChip array such that the ratio of "average" prostate cancer to "average" normal adult tissues was greater than or equal to 3.44. The "average" prostate cancer level was set to the 85th percentile amongst 73 prostate cancers. The "average" normal adult tissue level was set to the 85th percentile amongst 162 non-malignant tissues. In order to remove gene-specific background levels of non-specific hybridization, the 7.5th percentile value amongst the 162 non-malignant tissues was subtracted from both the numerator and the denominator before the ratio was evaluated.

Pkey: Unique Eos probeset identifier number
ExAccn: Exemplar Accession number, Genbank accession number
UnigeneID: Unigene number
Unigene Title: Unigene gene title
R1: Ratio of tumor to normal tissue

Pkey	ExAccn	UnigeneID	Unigene Title	R1
446057	AI420227	Hs.149358	ESTs, Weakly similar to A46010 X-linked	86.42
400302	N48056	Hs.1915	folate hydrolase (prostate-specific memb	66.46
414569	AF109298	Hs.118258	prostate cancer associated protein 1	58.36
417407	AA923278	Hs.290905	ESTs, Weakly similar to protease [H.sapi	56.16
431579	AW971082	Hs.222886	ESTs, Weakly similar to TRHY_HUMAN TRICH	53.38
409361	NM_005982	Hs.54416	sine oculis homeobox (Drosophila) homolo	48.28
409731	AA125985	Hs.56145	thymosin, beta, identified in neuroblast	45.24
400298	AA032279	Hs.61635	six transmembrane epithelial antigen of	43.48
420154	AI093155	Hs.95420	JM27 protein	41.12
433466	AA508353	Hs.105314	relaxin 1 (H1)	39.88
400296	AA305627	Hs.139336	ATP-binding cassette, sub-family C (CFTR	38.42
400292	AA250737	Hs.72472	ESTs	38.00
432887	AI926047	Hs.162859	ESTs	36.48
439176	AI446444	Hs.190394	ESTs, Weakly similar to B28096 line-1 pr	36.45
430722	AW968543	Hs.203270	ESTs, Weakly similar to ALU1_HUMAN ALU S	33.20
437052	AA861697	Hs.120591	ESTs	33.02
418396	AI765805	Hs.26691	ESTs	32.68
434036	AI659131	Hs.197733	hypothetical protein MGC2849	32.44
407709	AA456135	Hs.23023	ESTs	32.10
426747	AA535210	Hs.171995	kallikrein 3, (prostate specific antigen	31.80
407168	R45175		ESTs	31.72
440260	AI972867	Hs.7130	copine IV	30.52
421513	X00949	Hs.105314	relaxin 1 (H1)	30.10
416370	N90470	Hs.203697	ESTs, Weakly similar to I38022 hypotheti	29.68
407122	H20276	Hs.31742	ESTs	29.24
400287	S39329	Hs.181350	kallikrein 2, prostatic	28.90
432244	AI669973	Hs.200574	ESTs	28.74
451939	U80456	Hs.27311	single-minded (Drosophila) homolog 2	28.74
415989	AI267700	Hs.111128	ESTs	28.34
418961	AW967646	Hs.23023	ESTs	27.34
425628	NM_004476	Hs.1915	folate hydrolase (prostate-specific memb	27.32
458509	AA654650	Hs.282906	ESTs	27.24
448290	AK002107	Hs.20843	Homo sapiens cDNA FLJ11245 fis, clone PL	27.16
428336	AA503115	Hs.183752	microseminoprotein, beta-	26.17
450096	AI682088	Hs.223368	holocarboxylase synthetase (biotin-[prop	25.60
400299	X07730	Hs.171995	kallikrein 3, (prostate specific antigen	24.91
437571	AA760894	Hs.153023	ESTs	24.74
453160	AI263307	Hs.146228	H2B histone family, member L	24.66
453096	AW294631	Hs.11325	ESTs	24.46
425075	AA506324	Hs.1852	acid phosphatase, prostate	24.23
407202	N58172	Hs.109370	ESTs	24.18

	424846	AU077324	Hs.1832	neuropeptide Y	23.57
	453370	AI470523	Hs.182356	ATP-binding cassette, sub-family C (CFTR	23.16
	422605	AA436989	Hs.121017	H2A histone family, member A	22.52
	444917	R68651	Hs.144997	ESTs	22.26
5	408826	AF216077	Hs.48376	Homo sapiens clone HB-2 mRNA sequence	22.02
	413597	AW302885	Hs.117183	ESTs	21.76
	426429	X73114	Hs.169849	myosin-binding protein C, slow-type	21.32
	435981	H74319	Hs.188620	ESTs	21.12
	432966	AA650114		ESTs	21.07
10	418848	AI820961	Hs.193465	ESTs	21.06
	405685				20.90
	443271	BE568568	Hs.195704	ESTs	19.98
	418819	AA228776	Hs.191721	ESTs	19.94
	420757	X78592	Hs.99915	androgen receptor (dihydrotestosterone r	19.72
15	418994	AA296520	Hs.89546	selectin E (endothelial adhesion molecul	19.56
	429918	AW873986	Hs.119383	ESTs	19.04
	415539	AI733881	Hs.72472	ESTs	18.43
	450382	AA397658	Hs.60257	Homo sapiens cDNA FLJ13598 fis, clone PL	18.34
	418829	AA516531	Hs.55999	NK homeobox (Drosophila), family 3, A	18.28
20	429984	AL050102	Hs.227209	hypothetical protein FLJ21617	17.82
	443822	AI087412	Hs.143611	ESTs, Weakly similar to 2004399A chromos	17.66
	431676	AI685464	Hs.292638	gb:tt88f04.x1 NCI_CGAP_Pr28 Homo sapiens	17.64
	410330	AW023630	Hs.46766	ESTs	17.52
	432441	AW292425	Hs.163484	ESTs	17.41
25	452792	AB037765	Hs.30652	KIAA1344 protein	17.39
	445472	AB006631	Hs.12784	Homo sapiens mRNA for KIAA0293 gene, par	17.00
	414565	AA502972	Hs.183390	hypothetical protein FLJ13590	16.82
	430487	D87742	Hs.241552	KIAA0268 protein	16.72
	431716	D89053	Hs.268012	fatty-acid-Coenzyme A ligase, long-chain	16.60
30	419536	AA603305		gb:np12d11.s1 NCI_CGAP_Pr3 Homo sapiens	16.50
	439677	R82331	Hs.164599	ESTs	16.46
	449625	NM_014253	Hs.23796	odz (odd Oz/ten-m, Drosophila) homolog 1	16.32
	408430	S79876	Hs.44926	dipeptidylpeptidase IV (CD26, adenosine	16.28
	447033	AI357412	Hs.157601	ESTs	16.02
35	453006	AI362575	Hs.167133	ESTs	15.74
	431474	AL133990	Hs.190642	ESTs	15.70
	420218	AW958037	Hs.22437	ribosomal protein L4	15.64
	408000	L11690	Hs.620	bullous pemphigoid antigen 1 (230/240kD)	15.54
	416208	AW291168	Hs.41295	ESTs, Weakly similar to MUC2_HUMAN MUCIN	15.48
40	430226	BE245562	Hs.2551	adrenergic, beta-2-, receptor, surface	15.40
	415263	AA948033	Hs.130853	ESTs	15.38
	432437	W07088	Hs.293685	ESTs	15.26
	428398	AI249368	Hs.98558	ESTs	15.21
	429900	AA460421	Hs.30875	ESTs	14.90
45	449156	AF103907	Hs.171353	prostate cancer antigen 3	14.89
	411096	U80034	Hs.68583	mitochondrial intermediate peptidase	14.81
	435974	U29680	Hs.37744	Homo sapiens beta-1 adrenergic receptor	14.76
	444484	AK002126	Hs.11260	hypothetical protein FLJ11264	14.76
	422728	AW937626	Hs.103262	ESTs, Weakly similar to ZN91_HUMAN ZINC	14.60
50	418601	AA279490	Hs.86368	calmegin	14.56
	448999	AF179274	Hs.22791	transmembrane protein with EGF-like and	14.55
	445885	AI734009	Hs.127699	KIAA1603 protein	14.44
	452712	AW638616		gb:RC5-LT0054-140200-013-D01 LT0054 Homo	14.22
	432189	AA527941		gb:nh30c04.s1 NCI_CGAP_Pr3 Homo sapiens	14.12
55	424565	AW102723	Hs.75295	guanylate cyclase 1, soluble, alpha 3	13.78
	429290	AF203032	Hs.198760	neurofilament, heavy polypeptide (200kD)	13.57
	419264	AA877104	Hs.293672	ESTs, Weakly similar to ALUB_HUMAN !!!!	13.40
	416445	AL043004	Hs.300678	KIAA0135 protein	13.32
	407275	AI364186		gb:qw34h07.x1 NCI_CGAP_Ut4 Homo sapiens	13.24
60	408369	R38438	Hs.182575	solute carrier family 15 (H+/peptide tra	13.21
	446720	AI439136	Hs.140546	ESTs	13.06
	434988	AI418055	Hs.161160	ESTs	13.02
	448172	N75276	Hs.135904	ESTs	12.98
	416182	NM_004354	Hs.79069	cyclin G2	12.94
65	420544	AA677577	Hs.98732	Homo sapiens Chromosome 16 BAC clone CIT	12.79
	445413	AA151342	Hs.12677	CGI-147 protein	12.64
	452588	AA889120	Hs.110637	homeo box A10	12.62
	407819	R42185	Hs.274803	ESTs	12.60
	433444	AW975324	Hs.129816	ESTs	12.60

	421059	AI654133	Hs.30212	thyroid receptor interacting protein 15	12.30
	420077	AW512260	Hs.87767	ESTs	12.24
	453930	AA419466	Hs.36727	hypothetical protein FLJ10903	12.22
	441610	AW576148	Hs.148376	ESTs	12.20
5	451009	AA013140	Hs.115707	ESTs	12.18
	433764	AW753676	Hs.39982	ESTs	12.16
	440286	U29589	Hs.7138	cholinergic receptor, muscarinic 3	12.04
	443912	R37257	Hs.184780	ESTs	11.92
	419526	AI821895	Hs.193481	ESTs	11.91
10	423073	BE252922	Hs.123119	MAD (mothers against decapentaplegic, Dr	11.87
	452784	BE463857	Hs.151258	hypothetical protein FLJ21062	11.86
	414422	AA147224	Hs.71814	ESTs	11.76
	450203	AF097994	Hs.301528	L-kynurenine/alpha-aminoadipate aminotra	11.68
	436679	AI127483	Hs.120451	ESTs, Weakly similar to unnamed protein	11.60
15	440901	AA909358	Hs.128612	ESTs	11.60
	448045	AJ297436	Hs.20166	prostate stem cell antigen	11.51
	433867	AW204232	Hs.279522	ESTs	11.50
	434980	AW770553	Hs.293640	sterol O-acyltransferase (acyl-Coenzyme	11.38
	425905	AB032959	Hs.161700	novel C3HC4 type Zinc finger (ring finge	11.33
20	434680	T11738	Hs.127574	ESTs	11.32
	449650	AF055575	Hs.297647	calcium channel, voltage-dependent, L ty	11.18
	431173	AW971198	Hs.294068	ESTs	11.16
	434539	AW748078	Hs.214410	ESTs, Weakly similar to MUC2_HUMAN MUCIN	11.16
	410037	AB020725	Hs.58009	KIAA0918 protein	11.14
25	417708	N74392	Hs.50495	ESTs	11.14
	458332	AI000341	Hs.220491	ESTs	11.12
	420381	D50640	Hs.301782	phosphodiesterase 3B, cGMP-inhibited	11.10
	425665	AK001050	Hs.159066	hypothetical protein FLJ10188	11.08
	425710	AF030880	Hs.159275	solute carrier family, member 4	11.08
30	428728	NM_016625	Hs.191381	hypothetical protein	11.04
	407021	U52077		gb:Human mariner1 transposase gene, comp	11.02
	410733	D84284	Hs.66052	CD38 antigen (p45)	11.02
	401714				10.90
	434485	AI623511	Hs.118567	ESTs	10.89
35	415786	AW419196	Hs.257924	hypothetical protein FLJ13782	10.87
	452340	NM_002202	Hs.505	ISL1 transcription factor, LIM/homeodoma	10.85
	453628	AW243307	Hs.170187	hypothetical protein	10.72
	408063	BE086548	Hs.42346	calcineurin-binding protein calsarcin-1	10.67
	417687	AI828596	Hs.250691	ESTs	10.64
40	434666	AF151103	Hs.112259	T cell receptor gamma locus	10.53
	432374	W68815	Hs.301885	Homo sapiens cDNA FLJ11346 fis, clone PL	10.50
	428819	AL135623	Hs.193914	KIAA0575 gene product	10.48
	413409	AI638418	Hs.21745	DEAD/H (Asp-Glu-Ala-Asp/His) box polypep	10.44
	428775	AA434579	Hs.143691	ESTs	10.21
45	436556	AI364997	Hs.7572	ESTs	10.20
	441690	R81733	Hs.33106	ESTs	10.14
	419852	AW503756	Hs.286184	hypothetical protein dJ551D2.5	10.10
	421991	NM_014918	Hs.110488	KIAA0990 protein	10.04
	423698	AA329796	Hs.1098	DKFZp434J1813 protein	10.02
50	452039	AI922988	Hs.172510	ESTs	10.00
	433043	W57554	Hs.125019	ESTs	9.98
	433927	AI557019	Hs.116467	small nuclear protein PRAC	9.97
	445424	AB028945	Hs.12696	cortactin SH3 domain-binding protein	9.96
	432240	AI694767	Hs.129179	Homo sapiens cDNA FLJ13581 fis, clone PL	9.88
55	433104	AL043002	Hs.128246	ESTs, Moderately similar to unnamed prot	9.84
	452744	AI267652	Hs.30504	Homo sapiens mRNA; cDNA DKFZp434E082 (fr	9.82
	431217	NM_013427	Hs.250830	Rho GTPase activating protein 6	9.75
	427398	AW390020	Hs.20415	chromosome 21 open reading frame 11	9.70
	446896	T15767	Hs.22452	Homo sapiens mRNA for KIAA1737 protein,	9.70
60	421470	R27496	Hs.1378	annexin A3	9.64
	406554				9.60
	401424				9.58
	407902	AL117474	Hs.41181	Homo sapiens mRNA; cDNA DKFZp727C191 (fr	9.56
	423545	AP000692	Hs.129781	chromosome 21 open reading frame 5	9.54
65	439024	R96696	Hs.35598	ESTs	9.51
	431548	AI834273	Hs.9711	novel protein	9.48
	409262	AK000631	Hs.52256	hypothetical protein FLJ20624	9.45
	446271	D82484	Hs.100469	ESTs	9.42
	448692	AW013907	Hs.224276	methylcrotonoyl-Coenzyme A carboxylase 2	9.26

	414140	AA281279	Hs.23317	hypothetical protein FLJ14681	9.24
	435980	AF274571	Hs.129142	deoxyribonuclease II beta	9.24
	421246	AW582962	Hs.300961	CGI-47 protein	9.20
	427304	AA761526	Hs.163853	ESTs	9.16
5	442914	AW188551	Hs.99519	hypothetical protein FLJ14007	9.16
	413627	BE182082	Hs.246973	ESTs	9.14
	439699	AF086534	Hs.187561	ESTs, Moderately similar to ALU1_HUMAN A	9.10
	437718	AI927288	Hs.196779	ESTs	9.07
	439820	AL360204	Hs.283853	Homo sapiens mRNA full length insert cDN	9.06
10	447342	AI199268	Hs.19322	Homo sapiens, Similar to RIKEN cDNA 2010	9.05
	446223	BE300091	Hs.119699	hypothetical protein FLJ12969	9.04
	410001	AB041036	Hs.57771	kallikrein 11	9.03
	424012	AW368377	Hs.137569	tumor protein 63 kDa with strong homolog	9.03
	441791	AW372449	Hs.175982	hypothetical protein FLJ21159	9.02
15	448206	BE622585	Hs.3731	ESTs, Moderately similar to I38022 hypot	9.02
	414269	AA298489		olfactory receptor, family 51, subfamily	8.99
	442081	AA401863	Hs.22380	ESTs	8.98
	420092	AA814043	Hs.88045	ESTs	8.85
	411630	U42349	Hs.71119	Putative prostate cancer tumor suppresso	8.80
20	421863	AI952677	Hs.108972	Homo sapiens mRNA; cDNA DKFZp434P228 (fr	8.80
	454141	AW138413	Hs.182356	ATP-binding cassette, sub-family C (CFTR	8.80
	418278	AI088489	Hs.83937	hypothetical protein	8.78
	428330	L22524	Hs.2256	matrix metalloproteinase 7 (matrilysin,	8.76
	432415	T16971	Hs.289014	ESTs, Weakly similar to A43932 mucin 2 p	8.75
25	424906	AI566086	Hs.153716	Homo sapiens mRNA for Hmob33 protein, 3'	8.74
	415245	N59650	Hs.27252	ESTs	8.72
	442409	BE208843	Hs.129544	hypothetical protein MGC15438	8.70
	404571				8.66
	418033	W68180	Hs.259855	elongation factor-2 kinase	8.64
30	456497	AW967956	Hs.123648	ESTs, Weakly similar to AF108460 1 ubinu	8.56
	405876				8.54
	448807	AI571940	Hs.7549	ESTs	8.52
	445372	N36417	Hs.144928	ESTs	8.48
	425171	AW732240	Hs.300615	ESTs	8.44
35	419968	X04430	Hs.93913	interleukin 6 (interferon, beta 2)	8.36
	407385	AA610150	Hs.272072	ESTs, Weakly similar to I38022 hypotheti	8.31
	433172	AB037841	Hs.102652	hypothetical protein ASH1	8.30
	422631	BE218919	Hs.118793	hypothetical protein FLJ10688	8.27
	412719	AW016610	Hs.129911	ESTs	8.24
40	418849	AW474547	Hs.53565	Homo sapiens PIG-M mRNA for mannosyltran	8.22
	444922	AI921750	Hs.144871	Homo sapiens cDNA FLJ13752 fis, clone PL	8.22
	427674	NM_003528	Hs.2178	H2B histone family, member Q	8.20
	432101	AI918950	Hs.11092	EphA3	8.17
	416288	H51299		gb:yp07c06.s1 Soares breast 3NbHBst Homo	8.15
45	404915				8.08
	440106	AA864968	Hs.127699	KIAA1603 protein	8.07
	442861	AA243837	Hs.57787	ESTs	8.06
	452259	AA317439	Hs.28707	signal sequence receptor, gamma (transio	8.06
	443250	AI041530	Hs.132107	ESTs	8.06
50	437267	AW511443	Hs.258110	ESTs	8.04
	452891	N75582	Hs.212875	ESTs, Weakly similar to DYH9_HUMAN CILI	8.02
	422219	AW978073		regulator of mitotic spindle assembly 1	8.00
	453049	BE537217	Hs.30343	ESTs	8.00
	439731	AI953135	Hs.45140	hypothetical protein FLJ14084	7.98
55	408554	AA836381	Hs.7323	nuclear receptor co-repressor/HDAC3 comp	7.94
	421154	AA284333	Hs.287631	Homo sapiens cDNA FLJ14269 fis, clone PL	7.94
	430107	AA465293	Hs.105069	ESTs	7.94
	433404	T32982	Hs.102720	ESTs	7.93
	450813	AI739625	Hs.203376	ESTs	7.90
60	416239	AL038450	Hs.48948	ESTs	7.85
	448212	AI475858		gb:tc87d07.x1 NCI_CGAP_CLL1 Homo sapiens	7.82
	449532	W74653	Hs.271593	ESTs, Moderately similar to A47582 B-cel	7.82
	413930	M86153	Hs.75618	RAB11A, member RAS oncogene family	7.80
	456191	AI420611	Hs.127832	ESTs	7.80
65	444858	AI199738	Hs.208275	ESTs, Weakly similar to ALUA_HUMAN !!!!	7.78
	457498	AI732230	Hs.191737	ESTs	7.78
	407235	D20569	Hs.169407	SAC2 (suppressor of actin mutations 2, y	7.76
	433759	AA680003	Hs.109363	Homo sapiens cDNA: FLJ23603 fis, clone L	7.74
	433805	AA706910	Hs.112742	ESTs	7.74

	426485	NM_006207	Hs.170040	platelet-derived growth factor receptor-	7.72
	446028	R44714	Hs.106795	Homo sapiens cDNA FLJ13136 fis, clone NT	7.72
	418555	AI417215	Hs.87159	hypothetical protein FLJ12577	7.70
	447499	AW262580	Hs.147674	protocadherin beta 16	7.70
5	419839	U24577	Hs.93304	phospholipase A2, group VII (platelet-ac	7.68
	416857	AA188775	Hs.292453	ESTs	7.68
	413801	M62246	Hs.35406	ESTs, Highly similar to unnamed protein	7.66
	425480	AB023198	Hs.158135	KIAA0981 protein	7.66
	420120	AL049610	Hs.95243	transcription elongation factor A (SII)-	7.64
10	424099	AF071202	Hs.139336	ATP-binding cassette, sub-family C (CFTR	7.64
	446307	T50083	Hs.9094	ESTs	7.63
	429220	AW207206	Hs.136319	ESTs	7.59
	420345	AW295230	Hs.25231	ESTs	7.54
	429208	AA447990	Hs.190478	ESTs	7.54
15	447247	AW369351	Hs.287955	Homo sapiens cDNA FLJ13090 fis, clone NT	7.53
	440995	T57773	Hs.10263	ESTs	7.53
	448706	AW291095	Hs.21814	interleukin 20 receptor, alpha	7.52
	410227	AB009284	Hs.61152	exostoses (multiple)-like 2	7.49
	431616	AA508552	Hs.195839	ESTs, Weakly similar to I38022 hypotheti	7.46
20	434217	AW014795	Hs.23349	ESTs	7.44
	431467	N71831	Hs.256398	Homo sapiens mRNA; cDNA DKFZp434E0528 (f	7.42
	448519	AW175665	Hs.244334	Homo sapiens prostein mRNA, complete cds	7.42
	446791	AI632278	Hs.34981	ESTs	7.40
	419743	AW408762	Hs.127478	Homo sapiens clone 24416 mRNA sequence	7.39
25	445855	BE247129	Hs.145569	ESTs	7.36
	425211	M18667	Hs.1867	progastricsin (pepsinogen C)	7.35
	419131	AA406293	Hs.301622	ESTs	7.34
	400294	N95796	Hs.179809	Homo sapiens prostein mRNA, complete cds	7.33
	441736	AW292779	Hs.169799	ESTs	7.28
30	427701	AA411101	Hs.221750	nuclear autoantigenic sperm protein (his	7.24
	457733	AW974812	Hs.291971	ESTs	7.24
	418432	M14156	Hs.85112	insulin-like growth factor 1 (somatomedi	7.22
	441201	AW118822	Hs.128757	ESTs	7.21
	419953	BE267154	Hs.125752	ESTs	7.20
35	419991	AJ000098	Hs.94210	eyes absent (Drosophila) homolog 1	7.20
	425018	BE245277	Hs.154196	E4F transcription factor 1	7.20
	424560	AA158727	Hs.150555	protein predicted by clone 23733	7.18
	435380	AA679001	Hs.192221	ESTs	7.14
	420658	AW965215	Hs.130707	ESTs	7.12
40	408291	AB023191	Hs.44131	KIAA0974 protein	7.10
	409110	AA191493	Hs.48778	niban protein	7.10
	414485	W27026	Hs.182625	VAMP (vesicle-associated membrane protei	7.10
	430039	BE253012	Hs.153400	ESTs, Weakly similar to ALU1_HUMAN ALU S	7.10
	450832	AW970602	Hs.105421	ESTs	7.10
45	417153	X57010	Hs.81343	collagen, type II, alpha 1 (primary oste	7.08
	412446	AI768015	Hs.92127	ESTs	7.07
	412953	Z45794	Hs.238809	ESTs	7.06
	418051	AW192535	Hs.19479	ESTs	7.06
	421566	NM_000399	Hs.1395	early growth response 2 (Krox-20 (Drosop	7.04
50	446999	AA151520	Hs.279525	hypothetical protein MGC4485	7.04
	440529	AW207640	Hs.16478	Homo sapiens cDNA: FLJ21718 fis, clone C	7.04
	441111	AI806867	Hs.126594	ESTs	7.01
	451027	AW519204	Hs.40808	ESTs	7.00
	408432	AW195262		gb:xn67b05.x1 NCI_CGAP_CML1 Homo sapiens	7.00
55	432223	AA333283	Hs.285336	Homo sapiens, clone IMAGE:3460280, mRNA	7.00
	444805	AB007899	Hs.12017	homolog of yeast ubiquitin-protein ligas	6.99
	414212	AA136569	Hs.295940	KIAA0187 gene product	6.98
	431725	X65724	Hs.2839	Norrie disease (pseudoglioma)	6.98
	449685	AW296669	Hs.66095	ESTs	6.97
60	447313	U92981	Hs.18081	Homo sapiens clone DT1P1B6 mRNA, CAG rep	6.96
	424590	AW966399	Hs.46821	hypothetical protein FLJ20086	6.94
	449655	AI021987	Hs.59970	ESTs	6.92
	419563	AA526235	Hs.193162	Homo sapiens cDNA FLJ11983 fis, clone HE	6.90
	434163	AW974720	Hs.25206	group XII secreted phospholipase A2	6.89
65	415809	Z32789	Hs.46601	ESTs	6.86
	425782	U66468	Hs.159525	cell growth regulatory with EF-hand doma	6.85
	417958	AA767382	Hs.193417	ESTs	6.84
	427408	AA583206	Hs.2156	RAR-related orphan receptor A	6.79
	445873	AA250970	Hs.251946	poly(A)-binding protein, cytoplasmic 1-l	6.74

	410718	AI920783	Hs.191435	ESTs	6.74
	432363	AA534489		gb:nf76g11.s1 NCI_CGAP_Co3 Homo sapiens	6.74
	436521	AW203986	Hs.213003	ESTs	6.73
	435604	AA625279	Hs.26892	uncharacterized bone marrow protein BM04	6.73
5	419083	AI479560	Hs.98613	Homo sapiens cDNA FLJ12292 fis, clone MA	6.72
	418245	AA088767	Hs.83883	transmembrane, prostate androgen induced	6.70
	420714	BE172704	Hs.222746	KIAA1610 protein	6.70
	412707	AW206373	Hs.16443	Homo sapiens cDNA: FLJ21721 fis, clone C	6.67
	421896	N62293	Hs.45107	ESTs	6.66
10	411078	AI222020	Hs.182364	CocoaCrisp	6.66
	452465	AA610211	Hs.34244	ESTs	6.66
	422763	AA033699	Hs.83938	ESTs, Moderately similar to MAS2_HUMAN M	6.66
	444618	AV653785	Hs.300171	ELL-RELATED RNA POLYMERASE II, ELONGATIO	6.64
	450164	AI239923	Hs.30098	ESTs	6.63
15	431060	AF039307	Hs.249171	homeo box A11	6.62
	408031	AA081395	Hs.42173	Homo sapiens cDNA FLJ10366 fis, clone NT	6.62
	420285	AA258124	Hs.293878	ESTs, Moderately similar to ZN91_HUMAN Z	6.62
	444670	H58373	Hs.37494	hypothetical protein MGC5370	6.62
	444489	AI151010	Hs.157774	ESTs	6.60
20	445685	AW779829	Hs.263436	gb:hn88a05.x1 NCI_CGAP_Kid11 Homo sapien	6.60
	435677	AA694142	Hs.293726	ESTs, Weakly similar to TSGA RAT TESTIS	6.59
	452221	C21322	Hs.11577	hypothetical protein FLJ22242	6.59
	431510	AA580082	Hs.112264	ESTs	6.56
	415874	AF091622	Hs.78893	KIAA0244 protein	6.54
25	418405	AI868282	Hs.11898	ESTs, Highly similar to KIAA1370 protein	6.54
	452768	AW069459	Hs.61539	ESTs	6.54
	401451				6.52
	416289	W26333		ESTs	6.52
	431778	AL080276	Hs.268562	regulator of G-protein signalling 17	6.51
30	409089	NM_014781	Hs.50421	KIAA0203 gene product	6.50
	442833	AA328153	Hs.88201	ESTs, Weakly similar to A Chain A, Cryst	6.50
	431992	NM_002742	Hs.2891	protein kinase C, mu	6.49
	418833	AW974899	Hs.292776	ESTs	6.48
	429163	AA884766		gb:am20a10.s1 Soares_NFL_T_GBC_S1 Homo s	6.46
35	430403	AF039390	Hs.241382	tumor necrosis factor (ligand) superfami	6.46
	443058	AW451642	Hs.16732	ESTs	6.46
	418564	AA631143	Hs.179809	Homo sapiens prostein mRNA, complete cds	6.44
	432674	AA641092	Hs.257339	ESTs, Weakly similar to I38022 hypotheti	6.44
	423600	AI633559	Hs.29076	ESTs	6.44
40	404253				6.42
	433610	AA806822	Hs.112547	ESTs	6.42
	421552	AF026692	Hs.105700	secreted frizzled-related protein 4	6.41
	407118	AA156790	Hs.262036	ESTs, Weakly similar to Z223_HUMAN ZINC	6.40
	408608	N79738	Hs.136102	KIAA0853 protein	6.40
45	421452	AI925946	Hs.104530	fetal hypothetical protein	6.40
	433285	AW975944	Hs.237396	ESTs	6.40
	434926	BE543269	Hs.50252	mitochondrial ribosomal protein L32	6.40
	446189	H85224	Hs.214013	ESTs	6.40
	416806	NM_000288	Hs.79993	peroxisomal biogenesis factor 7	6.38
50	416467	H57585	Hs.37467	ESTs	6.36
	453403	BE466639	Hs.61779	Homo sapiens cDNA FLJ13591 fis, clone PL	6.34
	429769	NM_004917	Hs.218366	kallikrein 4 (prostase, enamel matrix, p	6.34
	423642	AW452650	Hs.157148	hypothetical protein MGC13204	6.32
	425843	BE313280	Hs.159627	death associated protein 3	6.32
55	439221	AA737106	Hs.32250	ESTs, Moderately similar to I78885 serin	6.32
	428194	AA765603	Hs.180877	H3 histone, family 3B (H3.3B)	6.30
	431958	X63629	Hs.2877	cadherin 3, type 1, P-cadherin (placenta	6.30
	439366	AF100143	Hs.6540	fibroblast growth factor 13	6.30
	452789	AW081626	Hs.242561	ESTs	6.30
60	416836	D54745	Hs.80247	cholecystokinin	6.30
	436962	AW377314	Hs.5364	DKFZP564I052 protein	6.29
	433383	AF034837	Hs.192731	double-stranded RNA specific adenosine d	6.29
	418636	AW749855		gb:QV4-BT0534-281299-053-c05 BT0534 Homo	6.26
	450728	AW162923	Hs.25363	presenilin 2 (Alzheimer disease 4)	6.25
65	440293	AI004193	Hs.22123	ESTs	6.24
	453745	AA952989	Hs.63908	hypothetical protein MGC14726	6.24
	426595	AW971980	Hs.62402	p21/Cdc42/Rac1-activated kinase 1 (yeast	6.24
	444412	AI147652	Hs.216381	Homo sapiens clone HH409 unknown mRNA	6.24
	413384	NM_000401	Hs.75334	exostoses (multiple) 2	6.22

	426320	W47595	Hs.169300	transforming growth factor, beta 2	6.22
	423349	AF010258	Hs.127428	homeo box A9	6.20
	429165	AW009886	Hs.118258	prostate cancer associated protein 1	6.18
	424800	AL035588	Hs.153203	MyoD family inhibitor	6.18
5	409564	AA045857	Hs.54943	fracture callus 1 (rat) homolog	6.16
	438796	W67821	Hs.109590	genethonin 1	6.16
	425451	AF242769	Hs.157461	mesenchymal stem cell protein DSC54	6.14
	451663	AI872360	Hs.209293	ESTs	6.14
	413623	AA625721	Hs.246973	ESTs	6.12
10	452232	AW020603	Hs.271698	radial spoke protein 3	6.12
	453390	AA862496	Hs.28482	ESTs	6.12
	435542	AA687376	Hs.269533	ESTs	6.12
	420424	AB033036	Hs.97594	KIAA1210 protein	6.11
	407103	AA424881	Hs.256301	hypothetical protein MGC13170	6.10
15	409734	BE161664	Hs.56155	hypothetical protein	6.10
	432686	BE223007	Hs.152460	Homo sapiens cDNA FLJ12909 fis, clone NT	6.10
	438361	AA805666	Hs.146217	Homo sapiens cDNA: FLJ23077 fis, clone L	6.10
	411479	AW848047		gb:IL3-CT0214-291299-052-A12 CT0214 Homo	6.10
	438849	W28948	Hs.10762	ESTs	6.08
20	452726	AF188527	Hs.61661	ESTs, Weakly similar to AF174605 1 F-box	6.08
	445895	D29954	Hs.13421	KIAA0056 protein	6.08
	440774	AI420611	Hs.127832	ESTs	6.07
	422583	AA410506	Hs.118578	KIAA0874 protein	6.06
	427500	AW970017	Hs.293948	ESTs, Weakly similar to S65657 alpha-1C-	6.04
25	443646	AI085198	Hs.298699	ESTs	6.04
	410566	AA373210	Hs.43047	Homo sapiens cDNA FLJ13585 fis, clone PL	6.02
	417845	AL117461	Hs.82719	Homo sapiens mRNA; cDNA DKFZp586F1822 (f	6.02
	430273	AI311127	Hs.125522	ESTs	6.02
	434792	AA649253	Hs.132458	ESTs	6.01
30	442490	AW965078	Hs.30212	thyroid receptor interacting protein 15	6.01
	420026	AI831190	Hs.166676	ESTs	6.00
	437782	AI370876	Hs.123163	exportin 1 (CRM1, yeast, homolog)	6.00
	447359	NM_012093	Hs.18268	adenylate kinase 5	6.00
	447713	AI420733	Hs.207083	ESTs	6.00
35	451073	AI758905	Hs.206063	ESTs	6.00
	451640	AA195601	Hs.26771	Human DNA sequence from clone 747H23 on	6.00
	410889	X91662	Hs.66744	twist (Drosophila) homolog (acrocephalos	5.97
	441222	AI277237	Hs.44208	hypothetical protein FLJ23153	5.96
	447732	AI758398	Hs.161318	ESTs	5.96
40	437756	AA767537	Hs.197096	ESTs	5.95
	408829	NM_006042	Hs.48384	heparan sulfate (glucosamine) 3-O-sulfot	5.94
	453911	AW503857	Hs.4007	Sarcolemmal-associated protein	5.94
	414085	AA114016	Hs.75746	aldehyde dehydrogenase 1 family, member	5.93
	408875	NM_015434	Hs.48604	DKFZP434B168 protein	5.92
45	439451	AF086270	Hs.278554	heterochromatin-like protein 1	5.92
	423853	AB011537	Hs.133466	slit (Drosophila) homolog 1	5.91
	453060	AW294092	Hs.21594	hypothetical protein MGC15754	5.91
	420407	AA814732	Hs.145010	lipopolysaccaride-specific response 5-li	5.91
	450480	X82125	Hs.25040	zinc finger protein 239	5.90
50	408446	AW450669	Hs.45068	hypothetical protein DKFZp434I143	5.88
	421039	NM_003478	Hs.101299	cullin 5	5.88
	451684	AF216751	Hs.26813	CDA14	5.88
	436063	AK000028	Hs.250867	ribosomal protein S24	5.86
	410507	AA355288	Hs.271408	transitional epithelia response protein	5.86
55	420179	N74530	Hs.21168	ESTs	5.84
	453878	AW964440	Hs.19025	DC32	5.84
	452270	AW975014	Hs.26	ferrochelatase (protoporphyria)	5.83
	435867	AA954229	Hs.114052	ESTs	5.82
	417683	AW566008	Hs.239154	ankyrin repeat, family A (RFXANK-like),	5.82
60	432005	AA524190	Hs.120777	ESTs, Weakly similar to ELL2_HUMAN RNA P	5.81
	406815	AA833930	Hs.288036	tRNA isopentenylpyrophosphate transferas	5.80
	437980	R50393	Hs.278436	KIAA1474 protein	5.80
	425856	AA364908	Hs.99927	hypothetical protein FLJ13993	5.79
	400301	X03635	Hs.1657	estrogen receptor 1	5.78
65	446261	AA313893	Hs.13399	hypothetical protein FLJ12615 similar to	5.78
	410141	R07775	Hs.287657	Homo sapiens cDNA: FLJ21291 fis, clone C	5.77
	427258	AA400091	Hs.39421	ESTs	5.76
	419108	AA389724	Hs.191264	ESTs, Weakly similar to ALU7_HUMAN ALU S	5.76
	442029	AW956698	Hs.14456	neural precursor cell expressed, develop	5.76

	407783	AW996872	Hs.172028	a disintegrin and metalloproteinase doma	5.75
	434408	AI031771	Hs.132586	ESTs	5.74
	415077	L41607	Hs.934	glucosaminyl (N-acetyl) transferase 2, I	5.74
	432435	BE218886	Hs.282070	ESTs	5.74
5	433313	W20128	Hs.296039	ESTs	5.73
	431740	N75450	Hs.183412	ESTs, Moderately similar to AF116721 67	5.73
	412991	AW949013		gb:QV4-FT0005-110500-201-e12 FT0005 Homo	5.72
	418852	BE537037	Hs.273294	hypothetical protein FLJ20069	5.72
	418882	NM_004996	Hs.89433	ATP-binding cassette, sub-family C (CFTR	5.72
10	446867	AB007891	Hs.16349	KIAA0431 protein	5.72
	437866	AA156781	Hs.83992	metallothionein 1E (functional)	5.72
	410232	AW372451	Hs.61184	CGI-79 protein	5.70
	414452	AA454038	Hs.29032	ESTs	5.70
	422762	AL031320	Hs.119976	Human DNA sequence from clone RP1-20N2 o	5.70
15	428730	AA625947	Hs.25750	ESTs	5.70
	431571	AW500486	Hs.180610	splicing factor proline/glutamine rich (	5.70
	433393	AF038564	Hs.98074	itchy (mouse homolog) E3 ubiquitin prote	5.70
	450616	AL133067	Hs.25214	hypothetical protein	5.70
	443774	AL117428	Hs.9740	DKFZP434A236 protein	5.69
20	446100	AW967109	Hs.13804	hypothetical protein dJ462O23.2	5.69
	419168	AI336132	Hs.33718	Homo sapiens cDNA FLJ12641 fis, clone NT	5.68
	416653	AA768553	Hs.77496	metallothionein 1E (functional)	5.67
	452679	Z42387	Hs.4299	transmembrane, prostate androgen induced	5.66
	450244	AA007534	Hs.125062	ESTs	5.66
25	408621	AI970672	Hs.46638	chromosome 11 open reading frame 8	5.65
	450325	AI935962	Hs.26269	ESTs	5.65
	439671	AW162840	Hs.6641	kinesin family member 5C	5.64
	452387	AI680772	Hs.4316	trinucleotide repeat containing 12	5.64
	413992	W26276	Hs.136075	RNA, U2 small nuclear	5.63
30	444151	AW972917	Hs.128749	alpha-methylacyl-CoA racemase	5.63
	417791	AW965339	Hs.111471	ESTs	5.62
	410196	AI936442	Hs.59838	hypothetical protein FLJ10808	5.60
	415123	D60925		ESTs	5.60
	429170	NM_001394	Hs.2359	dual specificity phosphatase 4	5.60
35	434415	BE177494		gb:RC6-HT0596-270300-011-C05 HT0596 Homo	5.60
	440738	AI004650	Hs.225674	WD repeat domain 9	5.60
	443830	AI142095	Hs.143273	ESTs	5.60
	449603	AI655662	Hs.197698	ESTs	5.60
	414342	AA742181	Hs.75912	KIAA0257 protein	5.59
40	422634	NM_016010	Hs.118821	CGI-62 protein	5.56
	435047	AA454985	Hs.54973	cadherin-like protein VR20	5.55
	400268				5.55
	452055	AI377431	Hs.293772	hypothetical protein MGC10858	5.54
	437073	AI885608	Hs.94122	ESTs	5.54
45	434072	H70854	Hs.283059	Homo sapiens PRO1082 mRNA, complete cds	5.53
	418339	AA639902	Hs.104215	ESTs, Moderately similar to SPCN_HUMAN S	5.52
	434551	BE387162	Hs.280858	ESTs, Highly similar to A35661 DNA excis	5.52
	439569	AW602166	Hs.222399	CEGP1 protein	5.51
	441102	AA973905	Hs.16003	intermediate filament protein syncoilin	5.50
50	448310	AI480316		gb:tm26h09.x1 Soares_NFL_T_GBC_S1 Homo s	5.50
	413173	BE076928	Hs.70980	ESTs	5.48
	436246	AW450963	Hs.119991	ESTs	5.48
	449300	AI656959	Hs.222165	ESTs	5.48
	452823	AB012124	Hs.30696	transcription factor-like 5 (basic helix	5.48
55	451403	AA885569	Hs.15727	Homo sapiens cDNA FLJ14511 fis, clone NT	5.46
	417061	AI675944	Hs.188691	Homo sapiens cDNA FLJ12033 fis, clone HE	5.44
	429126	AW172356	Hs.99083	ESTs	5.44
	431316	AA502663	Hs.145037	ESTs	5.44
	439192	AW970536	Hs.105413	ESTs	5.44
60	431938	AA938471	Hs.115242	specific granule protein (28 kDa); cyste	5.44
	451552	AA047233	Hs.33810	ESTs	5.43
	416991	N36389	Hs.295091	KIAA0226 gene product	5.42
	427638	AA406411	Hs.208341	ESTs, Weakly similar to KIAA0989 protein	5.42
	427718	AI798680	Hs.25933	ESTs	5.42
65	438710	AA833907	Hs.178724	ESTs, Weakly similar to ALU1_HUMAN ALU S	5.42
	406076	AL390179	Hs.137011	Homo sapiens mRNA; cDNA DKFZp547P134 (fr	5.40
	431263	AW129203	Hs.13743	ESTs	5.40
	421264	AL039123	Hs.103042	microtubule-associated protein 1B	5.38
	421685	AF189723	Hs.106778	ATPase, Ca++ transporting, type 2C, memb	5.37

	408460	AA054726	Hs.285574	ESTs	5.36
	409091	AW970386	Hs.269423	ESTs	5.36
	421987	AI133161	Hs.286131	CGI-101 protein	5.36
	428002	AA418703		gb:zv98c03.s1 Soares_NhHMPu_S1 Homo sapi	5.36
5	441217	AI922183	Hs.213246	ESTs	5.36
	426006	R49031	Hs.22627	ESTs	5.35
	422806	BE314767	Hs.1581	glutathione S-transferase theta 2	5.34
	432281	AK001239	Hs.274263	hypothetical protein FLJ10377	5.32
	451982	F13036	Hs.27373	Homo sapiens mRNA; cDNA DKFZp564O1763 (f	5.32
10	421129	BE439899	Hs.89271	ESTs	5.31
	444042	NM_004915	Hs.10237	ATP-binding cassette, sub-family G (WHIT	5.31
	410150	AW382942	Hs.6774	ESTs	5.30
	423952	AW877787	Hs.136102	KIAA0853 protein	5.30
	452822	X85689	Hs.288617	hypothetical protein FLJ22621	5.30
15	447752	M73700	Hs.347	lactotransferrin	5.29
	441766	R53790	Hs.23294	hypothetical protein FLJ14393	5.29
	431359	AW993522	Hs.292934	ESTs	5.27
	427212	AW293849	Hs.58279	ESTs, Weakly similar to ALU7_HUMAN ALU S	5.27
	449916	T60525	Hs.299221	pyruvate dehydrogenase kinase, isoenzyme	5.27
20	454014	AW016670	Hs.233275	ESTs	5.27
	419714	AA758751	Hs.98216	ESTs	5.26
	428845	AL157579	Hs.153610	KIAA0751 gene product	5.26
	417333	AL157545	Hs.42179	bromodomain and PHD finger containing, 3	5.24
	419986	AI345455	Hs.78915	GA-binding protein transcription factor,	5.24
25	407182	AA312551	Hs.230157	ESTs	5.22
	420111	AA255652		gb:zs21h11.r1 NCI_CGAP_GCB1 Homo sapiens	5.22
	428058	AI821625	Hs.191602	ESTs	5.22
	459551	AI472808		gb:tj70e07.x1 Soares_NSF_F8_9W_OT_PA_P_S	5.22
	432524	AI458020	Hs.293287	ESTs	5.22
30	436207	AA334774	Hs.12845	hypothetical protein MGC13159	5.22
	410870	U81599	Hs.66731	homeo box B13	5.22
	451418	BE387790	Hs.26369	hypothetical protein FLJ20287	5.22
	409757	NM_001898	Hs.123114	cystatin SN	5.21
	441124	T97717	Hs.119563	ESTs	5.21
35	428593	AW207440	Hs.185973	degenerative spermatocyte (homolog Droso	5.21
	436401	AI087958	Hs.29088	ESTs	5.20
	437113	AA744693		gb:ny26c10.s1 NCI_CGAP_GCB1 Homo sapiens	5.20
	450947	AI745400	Hs.204662	ESTs	5.20
	453279	AW893940	Hs.59698	ESTs	5.20
40	445467	AI239832	Hs.15617	ESTs, Weakly similar to ALU4_HUMAN ALU S	5.19
	448944	AB014605	Hs.22599	atrophin-1 interacting protein 1; activi	5.19
	412198	AA937111	Hs.69165	ESTs	5.18
	422646	H87863	Hs.151380	ESTs, Weakly similar to T16584 hypotheti	5.18
	438986	AF085888	Hs.269307	ESTs	5.18
45	453954	AW118336	Hs.75251	DEAD/H (Asp-Glu-Ala-Asp/His) box binding	5.18
	447541	AK000288	Hs.18800	hypothetical protein FLJ20281	5.18
	434029	AA621763	Hs.170434	Homo sapiens cDNA FLJ14242 fis, clone OV	5.16
	459294	AW977286	Hs.169531	RBP1-like protein	5.16
	429441	AJ224172	Hs.204096	lipophilin B (uteroglobin family member)	5.16
50	424692	AA429834	Hs.151791	KIAA0092 gene product	5.15
	427359	AW020782	Hs.79881	Homo sapiens cDNA: FLJ23006 fis, clone L	5.15
	419872	AI422951	Hs.146162	ESTs	5.15
	429422	AK001494	Hs.202596	Homo sapiens cDNA FLJ10632 fis, clone NT	5.14
	448902	Z45998	Hs.22543	Homo sapiens mRNA; cDNA DKFZp761I1912 (f	5.14
55	459055	N23235	Hs.30567	ESTs, Weakly similar to B34087 hypotheti	5.14
	431318	AA502700	Hs.293147	ESTs, Moderately similar to A46010 X-lin	5.14
	452953	AI932884	Hs.271741	ESTs, Weakly similar to A46010 X-linked	5.13
	428372	AK000684	Hs.183887	hypothetical protein FLJ22104	5.12
	434401	AI864131	Hs.71119	Putative prostate cancer tumor suppresso	5.12
60	416434	AW163045	Hs.79334	nuclear factor, interleukin 3 regulated	5.11
	410268	AA316181	Hs.61635	six transmembrane epithelial antigen of	5.10
	417517	AF001176	Hs.82238	POP4 (processing of precursor , S. cerev	5.10
	453616	NM_003462	Hs.33846	dynein, axonemal, light intermediate pol	5.10
	427958	AA418000	Hs.98280	potassium intermediate/small conductance	5.09
65	407945	X69208	Hs.606	ATPase, Cu++ transporting, alpha polypep	5.08
	425154	NM_001851	Hs.154850	collagen, type IX, alpha 1	5.08
	412863	AA121673	Hs.59757	zinc finger protein 281	5.06
	420807	AA280627	Hs.57846	ESTs	5.06
	430568	AA769221	Hs.270847	delta-tubulin	5.06

	433687	AA743991		gb:ny57g01.s1 NCI_CGAP_Pr18 Homo sapiens	5.06
	438375	AW015940	Hs.232234	ESTs	5.06
	418092	R45154	Hs.106604	ESTs	5.06
	418576	AW968159	Hs.289104	Alu-binding protein with zinc finger dom	5.05
5	413328	Y15723	Hs.75295	guanylate cyclase 1, soluble, alpha 3	5.04
	414271	AK000275	Hs.75871	protein kinase C binding protein 1	5.04
	432729	AK000292	Hs.278732	hypothetical protein FLJ20285	5.04
	433433	AI692623	Hs.121513	Homo sapiens clone Z'3-1 placenta expres	5.04
	439662	H97552	Hs.269060	ESTs	5.04
10	439743	AL389956	Hs.283858	Homo sapiens mRNA full length insert cDN	5.04
	417511	AL049176	Hs.82223	chordin-like	5.02
	437814	AI088192	Hs.135474	ESTs, Weakly similar to DDX9_HUMAN ATP-D	5.02
	426342	AF093419	Hs.169378	multiple PDZ domain protein	5.02
	429782	NM_005754	Hs.220689	Ras-GTPase-activating protein SH3-domain	5.02
15	429975	AI167145	Hs.165538	ESTs	5.02
	436209	AW850417	Hs.254020	ESTs, Moderately similar to unnamed prot	5.02
	438571	AW020775	Hs.56022	ESTs	5.02
	450223	AA418204	Hs.241493	natural killer-tumor recognition sequenc	5.02
	408267	AW380525	Hs.267705	tubulin-specific chaperone e	5.01
20	417730	Z44761		gb:HSC28F061 normalized infant brain cDN	5.00
	425465	L18964	Hs.1904	protein kinase C, iota	5.00
	430599	NM_004855	Hs.247118	phosphatidylinositol glycan, class B	5.00
	450961	AW978813	Hs.250867	metallothionein 1E (functional)	5.00
	451386	AB029006	Hs.26334	spastic paraplegia 4 (autosomal dominant	5.00
25	420380	AA640891	Hs.102406	ESTs	4.99
	424947	R77952	Hs.239625	ESTs, Weakly similar to alternatively sp	4.99
	442653	BE269247	Hs.170226	gb:601185486F1 NIH_MGC_8 Homo sapiens cD	4.98
	457211	AW972565	Hs.32399	ESTs, Weakly similar to S51797 vasodilat	4.97
	425851	NM_001490	Hs.159642	glucosaminyl (N-acetyl) transferase 1, c	4.97
30	446279	AA490770	Hs.182382	ESTs	4.96
	433377	AI752713	Hs.43845	ESTs	4.96
	450218	R02018	Hs.168640	ankylosis, progressive (mouse) homolog	4.96
	412715	NM_000947	Hs.74519	primase, polypeptide 2A (58kD)	4.94
	448164	R61680	Hs.26904	ESTs, Moderately similar to Z195_HUMAN Z	4.94
35	420121	AW968271	Hs.191534	ESTs, Weakly similar to ALU1_HUMAN ALU S	4.94
	421689	N87820	Hs.106826	KIAA1696 protein	4.93
	445808	AV655234	Hs.298083	ESTs, Moderately similar to PC4259 ferri	4.92
	416533	BE244053	Hs.79362	retinoblastoma-like 2 (p130)	4.92
	418049	AA211467	Hs.190488	Homo sapiens, Similar to nuclear localiz	4.92
40	436039	AW023323	Hs.121070	ESTs	4.92
	432653	N62096	Hs.293185	ESTs, Weakly similar to JC7328 amino aci	4.91
	420324	AF163474	Hs.96744	prostate androgen-regulated transcript 1	4.91
	403047				4.91
	436899	AA764852	Hs.291567	ESTs	4.90
45	431117	AF003522	Hs.250500	delta (Drosophila)-like 1	4.90
	427617	D42063	Hs.179825	RAN binding protein 2	4.88
	428804	AK000713	Hs.193736	*hypothetical protein FLJ20706*	4.88
	433050	AI093930	Hs.163440	Homo sapiens cDNA: FLJ21000 fis, clone C	4.88
	418575	AA225313	Hs.222886	ESTs, Weakly similar to TRHY_HUMAN TRICH	4.86
50	432615	AA557191	Hs.55028	ESTs, Weakly similar to I54374 gene NF2	4.86
	412652	AI801777	Hs.6774	ESTs	4.86
	432473	AI202703	Hs.152414	ESTs	4.86
	449071	NM_005872	Hs.22960	breast carcinoma amplified sequence 2	4.86
	450654	AJ245587	Hs.25275	Kruppel-type zinc finger protein	4.85
55	418866	T65754	Hs.100489	gb:yc11c07.s1 Stratagene lung (937210) H	4.85
	407596	R86913		gb:yq30f05.r1 Soares fetal liver spleen	4.84
	456516	BE172704	Hs.222746	KIAA1610 protein	4.84
	426501	AW043782	Hs.293616	ESTs	4.84
	448730	AB032983	Hs.21894	KIAA1157 protein	4.84
60	458339	AW976853	Hs.172843	ESTs	4.83
	422083	NM_001141	Hs.111256	arachidonate 15-lipoxygenase, second typ	4.82
	420159	AI572490	Hs.99785	Homo sapiens cDNA: FLJ21245 fis, clone C	4.82
	424103	NM_001918	Hs.139410	dihydrolipoamide branched chain transacy	4.82
	449535	W15267	Hs.23672	low density lipoprotein receptor-related	4.82
65	422048	NM_012445	Hs.288126	spondin 2, extracellular matrix protein	4.82
	416737	AF154335	Hs.79691	LIM domain protein	4.82
	419972	AL041465	Hs.294038	golgin-67	4.81
	420235	AA256756	Hs.31178	ESTs	4.81
	423412	AF109300	Hs.147924	prostate cancer associated protein 5	4.80

429598	AA811257	Hs.269710	ESTs	4.80
457114	AI821625	Hs.191602	ESTs	4.80
421828	AW891965	Hs.289109	histone deacetylase 3	4.79
424602	AK002055	Hs.301129	hypothetical protein FLJ11193	4.78
428364	AA426565	Hs.160541	ESTs, Moderately similar to ALU1_HUMAN A	4.78
452335	AW188944	Hs.61272	ESTs	4.78
410765	AI694972	Hs.66180	nucleosome assembly protein 1-like 2	4.77
421040	AA715026	Hs.135280	ESTs	4.76
421518	AI056392	Hs.208819	ESTs	4.76
452560	BE077084		ESTs	4.76
409752	AW963990		gb:EST376063 MAGE resequences, MAGH Homo	4.75
439703	AF086538	Hs.196245	ESTs	4.75
418836	AI655499	Hs.161712	ESTs	4.74
450642	R39773	Hs.7130	copine IV	4.74
419879	Z17805	Hs.93564	Homer, neuronal immediate early gene, 2	4.74
411440	AW749402		gb:QV4-BT0383-281299-061-c06 BT0383 Homo	4.74
450649	NM_001429	Hs.297722	E1A binding protein p300	4.74
408738	NM_014785	Hs.47313	KIAA0258 gene product	4.73
435020	AW505076	Hs.301855	DiGeorge syndrome critical region gene 8	4.72
411624	BE145964		KIAA0594 protein	4.72
439360	AA448488	Hs.55346	ribosomal protein L44	4.72
440491	R35252	Hs.24944	ESTs, Weakly similar to 2109260A B cell	4.72
442611	BE077155	Hs.177537	hypothetical protein DKFZp761B1514	4.72
443555	N71710	Hs.21398	ESTs, Moderately similar to A Chain A, H	4.72
453800	BE300741	Hs.288416	hypothetical protein FLJ13340	4.72
457528	AW973791	Hs.292784	ESTs	4.72
416795	AI497778	Hs.168053	HBV pX associated protein-8	4.71
407302	R74206	Hs.268755	ESTs, Weakly similar to I78885 serine/th	4.71
404721				4.70
426261	AW242243	Hs.168670	peroxisomal farnesylated protein	4.70
431924	AK000850	Hs.272203	Homo sapiens cDNA FLJ20843 fis, clone AD	4.70
435256	AF193766	Hs.13872	cytokine-like protein C17	4.70
438295	AI394151	Hs.37932	ESTs	4.70
442655	AW027457	Hs.30323	ESTs, Weakly similar to B34087 hypotheti	4.70
415788	AW628686	Hs.78851	KIAA0217 protein	4.69
442760	BE075297	Hs.10067	ESTs, Weakly similar to A43932 mucin 2 p	4.69
432432	AA541323	Hs.115831	ESTs	4.68
454398	AA463437	Hs.11556	Homo sapiens cDNA FLJ12566 fis, clone NT	4.68
452741	BE392914	Hs.30503	Homo sapiens cDNA FLJ11344 fis, clone PL	4.67
424853	BE549737	Hs.132967	Human EST clone 122687 mariner transposo	4.67
419706	C04649	Hs.77899	tropomyosin 1 (alpha)	4.66
412088	AI689496	Hs.108932	ESTs	4.65
416276	U41060	Hs.79136	LIV-1 protein, estrogen regulated	4.64
429281	AA830856	Hs.29808	Homo sapiens cDNA: FLJ21122 fis, clone C	4.64
448207	AI475490	Hs.170577	ESTs	4.64
408374	AW025430	Hs.155591	forkhead box F1	4.64
447162	BE328091	Hs.157396	ESTs, Weakly similar to A46010 X-linked	4.64
451900	AB023199	Hs.27207	KIAA0982 protein	4.63
421437	AW821252	Hs.104336	hypothetical protein	4.63
418624	AI734080	Hs.104211	ESTs	4.63
426172	AA371307	Hs.125056	ESTs	4.62
439831	AW136488	Hs.25545	ESTs	4.61
452994	AW962597	Hs.31305	KIAA1547 protein	4.61
457726	AI217477	Hs.194591	ESTs	4.60
434629	AA789081	Hs.4029	glioma-amplified sequence-41	4.60
403764				4.58
410659	AI080175	Hs.68826	ESTs	4.58
432383	AK000144	Hs.274449	Homo sapiens cDNA FLJ20137 fis, clone CO	4.58
451246	AW189232	Hs.39140	cutaneous T-cell lymphoma tumor antigen	4.58
433234	AB040928	Hs.65366	KIAA1495 protein	4.57
424983	AI742434	Hs.169911	ESTs	4.56
437812	AI582291	Hs.16846	ESTs, Weakly similar to O4HUD1 debrisoqu	4.56
438447	AI082883	Hs.167593	hypothetical protein FLJ13409; KIAA1711	4.55
434715	BE005346	Hs.116410	ESTs	4.55
447673	AI823987	Hs.182285	ESTs	4.54
408897	N50204	Hs.283709	lipopolysaccharide specific response-7 p	4.54
436645	AW023424	Hs.156520	ESTs	4.54
421247	BE391727	Hs.102910	general transcription factor IIH, polype	4.53
450377	AB033091	Hs.24936	KIAA1265 protein	4.53

	433644	AW342028	Hs.256112	gb:hb75d03.x1 NCI_CGAP_Ut2 Homo sapiens	4.53
	408321	AW405682	Hs.44205	cortistatin	4.53
	439225	AA192669	Hs.45032	ESTs	4.52
	440348	AW015802	Hs.47023	ESTs	4.52
5	446351	AW444551	Hs.258532	x 001 protein	4.52
	451212	AW902672	Hs.287334	ESTs	4.52
	430294	AI538226	Hs.135184	guanine nucleotide binding protein 4	4.52
	435005	U80743	Hs.4316	trinucleotide repeat containing 12	4.52
	448072	AI459306	Hs.24908	ESTs	4.50
10	403721				4.50
	451018	AW965599	Hs.247324	mitochondrial ribosomal protein S14	4.50
	453070	AK001465	Hs.31575	SEC63, endoplasmic reticulum translocon	4.49
	417412	X16896	Hs.82112	interleukin 1 receptor, type I	4.48
	439735	AI635386	Hs.142846	hypothetical protein	4.48
15	435663	AI023707	Hs.134273	ESTs	4.48
	424036	AA770688	Hs.81946	H2A histone family, member L	4.48
	426386	AA746850	Hs.174877	bladder cancer overexpressed protein	4.48
	408622	AA056060	Hs.202577	Homo sapiens cDNA FLJ12166 fis, clone MA	4.47
	444269	AI590346	Hs.146220	ESTs	4.47
20	430187	AI799909	Hs.158989	ESTs	4.46
	427761	AA412205	Hs.140996	ESTs	4.46
	430261	AA305127	Hs.237225	hypothetical protein HT023	4.46
	444169	AV648170	Hs.58756	ESTs	4.44
	430598	AK001764	Hs.247112	hypothetical protein FLJ10902	4.44
25	412903	BE007967	Hs.155795	ESTs	4.44
	417048	AI088775	Hs.55498	geranylgeranyl diphosphate synthase 1	4.44
	442710	AI015631	Hs.23210	ESTs	4.44
	457413	AA743462	Hs.165337	ESTs	4.44
	400303	AA242758	Hs.79136	LIV-1 protein, estrogen regulated	4.42
30	443268	AI800271	Hs.129445	hypothetical protein FLJ12496	4.42
	438209	AL120659	Hs.6111	aryl-hydrocarbon receptor nuclear transl	4.42
	431724	AA514535	Hs.283704	ESTs	4.41
	412280	AW205116	Hs.272814	hypothetical protein DKFZp434E1723	4.40
	440801	AA906366	Hs.190535	ESTs	4.40
35	452959	AI933416	Hs.189674	ESTs	4.40
	453861	AI026838	Hs.30120	ESTs, Weakly similar to NUCL_HUMAN NUCLE	4.40
	417421	AL138201	Hs.82120	nuclear receptor subfamily 4, group A, m	4.40
	447270	AC002551	Hs.331	general transcription factor IIIC, polyp	4.38
	433641	AF080229		gb:Human endogenous retrovirus K clone 1	4.38
40	447078	AW885727	Hs.301570	ESTs	4.38
	424242	AA337476		hypothetical protein MGC13102	4.37
	408170	AW204516	Hs.31835	ESTs	4.36
	448757	AI366784	Hs.48820	TATA box binding protein (TBP)-associate	4.36
	420021	AA252848	Hs.293557	ESTs	4.36
45	449694	AI659790	Hs.253302	ESTs	4.36
	453867	AI929383	Hs.108196	hypothetical protein DKFZp434N185	4.36
	458712	AI347502	Hs.173066	hypothetical protein FLJ20761	4.36
	417251	AW015242	Hs.99488	ESTs, Weakly similar to YK54_YEAST HYPOT	4.35
	434423	NM_006769	Hs.3844	LIM domain only 4	4.35
50	423427	AL137612	Hs.285848	KIAA1454 protein	4.34
	415715	F30364		ESTs	4.33
	404561				4.32
	422969	AA782536	Hs.122647	N-myristoyltransferase 2	4.32
	423685	BE350494	Hs.49753	uveal autoantigen with coiled coil domai	4.32
55	443977	AL120986	Hs.150627	ESTs, Weakly similar to I38022 hypotheti	4.32
	425071	NM_013989	Hs.154424	deiodinase, iodothyronine, type II	4.32
	431583	AL042613	Hs.262476	S-adenosylmethionine decarboxylase 1	4.31
	411379	AI816344	Hs.12554	ESTs, Weakly similar to NPL4_HUMAN NUCLE	4.30
	421476	AW953805	Hs.21887	ESTs	4.30
60	425178	H16097	Hs.161027	ESTs	4.30
	439262	AA832333	Hs.124399	ESTs	4.30
	442818	AK001741	Hs.8739	hypothetical protein FLJ10879	4.30
	421977	W94197	Hs.110165	ribosomal protein L26 homolog	4.29
	437114	AA836641	Hs.163085	ESTs	4.28
65	420195	N44348	Hs.300794	Homo sapiens cDNA FLJ11177 fis, clone PL	4.28
	418330	BE409405	Hs.94722	ESTs	4.27
	419750	AL079741	Hs.183114	Homo sapiens cDNA FLJ14236 fis, clone NT	4.26
	437065	AL036450	Hs.103238	ESTs	4.26
	455276	BE176479		gb:RC3-HT0585-160300-022-b09 HT0585 Homo	4.24

	416292	AA179233	Hs.42390	nasopharyngeal carcinoma susceptibility	4.24
	423740	Y07701	Hs.132243	aminopeptidase puromycin sensitive	4.24
	442023	AI187878	Hs.144549	ESTs	4.24
	426764	AA732524	Hs.151464	ESTs, Weakly similar to ALUC_HUMAN !!!!	4.23
5	454058	AI273419	Hs.135146	hypothetical protein FLJ13984	4.23
	456511	AA282330	Hs.145668	ESTs	4.22
	448330	AL036449	Hs.207163	ESTs	4.22
	424701	NM_005923	Hs.151988	mitogen-activated protein kinase kinase	4.21
	432621	AI298501	Hs.12807	ESTs, Weakly similar to T46428 hypotheti	4.20
10	445707	AI248720	Hs.114390	ESTs	4.20
	419910	AA662913	Hs.190173	ESTs, Weakly similar to A46010 X-linked	4.20
	424085	NM_002914	Hs.139226	replication factor C (activator 1) 2 (40	4.20
	440749	W22335	Hs.7392	hypothetical protein MGC3199	4.20
	442787	W93048	Hs.227203	hypothetical protein MGC2747	4.20
15	443414	R54594	Hs.25209	ESTs	4.20
	443556	AA256769	Hs.94949	methylmalonyl-CoA epimerase	4.20
	444170	AW613879	Hs.102408	ESTs	4.20
	446751	AA766998	Hs.85874	Human DNA sequence from clone RP11-16L21	4.20
	421041	N36914	Hs.14691	ESTs, Moderately similar to I38022 hypot	4.19
20	447476	BE293466	Hs.20880	ESTs, Weakly similar to I38022 hypotheti	4.19
	448543	AW897741	Hs.21380	Homo sapiens mRNA; cDNA DKFZp586P1124 (f	4.18
	410294	AB014515	Hs.288891	KIAA0615 gene product	4.18
	433607	AA602004	Hs.23260	ESTs	4.18
	435552	AI668636	Hs.193480	ESTs, Moderately similar to ALU6_HUMAN A	4.18
25	447124	AW976438	Hs.17428	RBP1-like protein	4.18
	453308	AW959731	Hs.32538	ESTs	4.17
	439328	W07411	Hs.118212	ESTs, Moderately similar to ALU3_HUMAN A	4.16
	430473	AW130690	Hs.299842	ESTs	4.16
	437257	AI283085	Hs.290931	ESTs, Weakly similar to YFJ7_YEAST HYPOT	4.16
30	438018	AK001160	Hs.5999	hypothetical protein FLJ10298	4.16
	443857	AI089292	Hs.287621	hypothetical protein FLJ14069	4.15
	446711	AF169692	Hs.12450	protocadherin 9	4.15
	419103	Z40229	Hs.96423	hypothetical protein FLJ23033	4.14
	405403				4.14
35	407378	AA299264		ESTs, Moderately similar to I38022 hypot	4.14
	408966	AW298602	Hs.197687	ESTs	4.14
	418727	AA227609	Hs.94834	ESTs	4.14
	434400	AI476211	Hs.186896	Homo sapiens cDNA FLJ11417 fis, clone HE	4.14
	438578	AA811244	Hs.164168	ESTs	4.14
40	450459	AI697193	Hs.299254	Homo sapiens cDNA: FLJ23597 fis, clone L	4.14
	429887	AW366286	Hs.145696	splicing factor (CC1.3)	4.13
	448148	NM_016578	Hs.20509	HBV pX associated protein-8	4.13
	450316	W84446	Hs.17850	hypothetical protein MGC4643	4.12
	417531	NM_003157	Hs.1087	serine/threonine kinase 2	4.12
45	431592	R69016	Hs.293871	hypothetical protein MGC10895s	4.12
	432463	AA548518	Hs.186733	ESTs	4.12
	433613	AA836126	Hs.5669	ESTs	4.12
	434739	AA804487	Hs.144130	ESTs	4.12
	438259	AW205969	Hs.131808	ESTs	4.12
50	425810	AI923627	Hs.31903	ESTs	4.10
	432672	AW973775	Hs.130760	myosin phosphatase, target subunit 2	4.10
	433345	AI681545	Hs.152982	hypothetical protein FLJ13117	4.10
	432712	AB016247	Hs.288031	sterol-C5-desaturase (fungal ERG3, delta	4.09
	453020	AL162039	Hs.31422	Homo sapiens mRNA; cDNA DKFZp434M229 (fr	4.09
55	412045	AA099802	Hs.4299	transmembrane, prostate androgen induced	4.09
	435114	AA775483	Hs.288936	mitochondrial ribosomal protein L9	4.08
	443204	AW205878	Hs.29643	Homo sapiens cDNA FLJ13103 fis, clone NT	4.08
	445459	AI476629	Hs.158465	likely ortholog of mouse putative IKK re	4.08
	436938	H46212	Hs.137221	ESTs	4.07
60	454119	BE549773	Hs.40510	uncoupling protein 4	4.06
	411000	N40449	Hs.201619	ESTs, Weakly similar to S38383 SEB4B pro	4.06
	418926	AA232658	Hs.87070	UDP-glucose:glycoprotein glucosyltransfe	4.06
	424432	AB037821	Hs.146858	protocadherin 10	4.06
	449673	AA002064	Hs.18920	ESTs	4.06
65	429299	AI620463	Hs.99197	hypothetical protein MGC13102	4.06
	422174	AL049325	Hs.112493	Homo sapiens mRNA; cDNA DKFZp564D036 (fr	4.05
	455497	AA112573	Hs.285691	Homo sapiens prostein mRNA, complete cds	4.05
	415138	C18356	Hs.78045	tissue factor pathway inhibitor 2	4.04
	402791				4.04

	426792	AL044854	Hs.172329	KIAA0576 protein	4.04
	438660	U95740	Hs.6349	Homo sapiens, clone IMAGE:3010666, mRNA,	4.04
	442768	AL048534	Hs.48458	ESTs, Weakly similar to ALU8_HUMAN ALU S	4.04
	447568	AF155655	Hs.18885	CGI-116 protein	4.04
5	428342	AI739168	Hs.131798	Homo sapiens cDNA FLJ13458 fis, clone PL	4.04
	453439	AI572438	Hs.32976	guanine nucleotide binding protein 4	4.02
	453857	AL080235	Hs.35861	DKFZP586E1621 protein	4.02
	428249	AA130914	Hs.183291	zinc finger protein 268	4.02
	432015	AL157504	Hs.159115	Homo sapiens mRNA; cDNA DKFZp586O0724 (f	4.02
10	445495	BE622641	Hs.38489	ESTs, Weakly similar to I38022 hypotheti	4.02
	451746	M86178		ESTs	4.02
	452211	AI985513	Hs.233420	ESTs	4.02
	453046	AA284040	Hs.219441	ESTs, Highly similar to CA5B_HUMAN CARBO	4.02
	456038	AA203285	Hs.294141	ESTs, Weakly similar to alternatively sp	4.02
15	452449	AW068658	Hs.20943	ESTs	4.02
	407204	R41933	Hs.140237	ESTs, Weakly similar to ALU1_HUMAN ALU S	4.01
	428046	AW812795	Hs.155381	ESTs, Moderately similar to I38022 hypot	4.01
	438520	AA706319	Hs.98416	ESTs	4.01
	443292	AK000213	Hs.9196	hypothetical protein	4.01
20	432715	AA247152	Hs.200483	ESTs, Weakly similar to KIAA1074 protein	4.00
	403797				4.00
	418347	AA216419	Hs.269295	gb:nc16e03.s1 NCI_CGAP_Pr1 Homo sapiens	4.00
	419459	AW291128	Hs.278422	DKFZP586G1122 protein	4.00
	420911	U77413	Hs.100293	O-linked N-acetylglucosamine (GlcNAc) tr	4.00
25	425176	AW015644	Hs.301430	TEA domain family member 1 (SV40 transcr	4.00
	447505	AL049266	Hs.18724	Homo sapiens mRNA; cDNA DKFZp564F093 (fr	4.00
	453773	AL133761		gb:DKFZp761C1413_r1 761 (synonym: hamy2)	4.00
	434384	AA631910	Hs.162849	ESTs	3.99
	422471	AA311027	Hs.271894	ESTs, Weakly similar to I38022 hypotheti	3.99
30	427386	AW836261	Hs.177486	ESTs	3.98
	433394	AI907753	Hs.93810	cerebral cavernous malformations 1	3.98
	441269	AW015206	Hs.178784	ESTs	3.97
	419629	AB020695	Hs.91662	KIAA0888 protein	3.96
	435008	AF150262	Hs.162898	ESTs	3.96
35	456649	R74441	Hs.117176	poly(A)-binding protein, nuclear 1	3.96
	418723	AA504428	Hs.10487	Homo sapiens, clone IMAGE:3954132, mRNA,	3.96
	428738	NM_000380	Hs.192803	xeroderma pigmentosum, complementation g	3.95
	430456	AA314998	Hs.241503	hypothetical protein	3.95
	422017	NM_003877	Hs.110776	STAT induced STAT inhibitor-2	3.95
40	409960	BE261944	Hs.153028	hexokinase 1	3.95
	455309	AW894017		gb:RC4-NN0027-150400-012-g04 NN0027 Homo	3.95
	450295	AI766732	Hs.201194	ESTs	3.94
	456660	AA909249	Hs.112282	solute carrier family 30 (zinc transport	3.94
	410908	AA121686	Hs.10592	ESTs	3.94
45	447145	AA761073	Hs.192943	TRAF family member-associated NFKB activ	3.94
	449318	AW236021	Hs.108788	Homo sapiens, Similar to RIKEN cDNA 5730	3.94
	449869	W57990	Hs.60059	Homo sapiens cDNA FLJ11478 fis, clone HE	3.94
	411887	AW182924	Hs.128790	ESTs	3.93
	437531	AI400752	Hs.112259	T cell receptor gamma locus	3.93
50	452238	F01811	Hs.187931	ESTs	3.93
	410486	AW235094	Hs.193424	zinc finger protein	3.92
	424882	AI379461	Hs.153636	far upstream element (FUSE) binding prot	3.92
	426269	H15302	Hs.168950	Homo sapiens mRNA; cDNA DKFZp566A1046 (f	3.92
	427043	AA397679	Hs.298460	ESTs	3.92
55	440404	AI015881	Hs.125616	mitochondrial ribosomal protein S5	3.92
	452762	AW501435	Hs.171409	v-akt murine thymoma viral oncogene homo	3.92
	453058	AW612293	Hs.288684	Homo sapiens cDNA FLJ11750 fis, clone HE	3.92
	423583	AL122055	Hs.129836	KIAA1028 protein	3.92
	408001	AA046458	Hs.95296	ESTs	3.92
60	419197	N48921	Hs.27441	KIAA1615 protein	3.91
	428695	AI355647	Hs.189999	purinergic receptor (family A group 5)	3.91
	401747				3.91
	410011	AB020641	Hs.57856	PFTAIRE protein kinase 1	3.91
	432205	AI806583	Hs.125291	ESTs	3.91
65	447857	AA081218	Hs.58608	Homo sapiens cDNA FLJ14206 fis, clone NT	3.91
	446494	AA463276	Hs.288906	WW Domain-Containing Gene	3.91
	409928	AL137163	Hs.57549	hypothetical protein dJ473B4	3.90
	411598	BE336654	Hs.70937	H3 histone family, member A	3.90
	424790	AL119344	Hs.13326	ESTs, Weakly similar to 2004399A chromos	3.90

	425707	AF115402	Hs.11713	E74-like factor 5 (ets domain transcript	3.90
	431325	AW026751	Hs.5794	ESTs, Weakly similar to 2109260A B cell	3.89
	451806	NM_003729	Hs.27076	RNA 3'-terminal phosphate cyclase	3.89
	401045				3.89
5	433023	AW864793	Hs.34161	thrombospondin 1	3.89
	452160	BE378541	Hs.279815	cysteine sulfinic acid decarboxylase-rel	3.89
	437372	AA323968	Hs.283631	hypothetical protein DKFZp547G183	3.89
	417067	AJ001417	Hs.81086	solute carrier family 22 (extraneuronal	3.88
	410467	AF102546	Hs.63931	dachshund (Drosophila) homolog	3.88
10	422660	AW297582	Hs.237062	hypothetical protein FLJ22548 similar to	3.88
	431930	AB035301	Hs.272211	cadherin 7, type 2	3.88
	453047	AW023798	Hs.286025	ESTs	3.88
	433891	AA613792		gb:no97h03.s1 NCI_CGAP_Pr2 Homo sapiens	3.88
	401785				3.88
15	431088	AA491824	Hs.196881	ESTs	3.88
	451952	AL120173	Hs.301663	ESTs	3.87
	422089	AA523172	Hs.103135	ESTs, Weakly similar to SFR4_HUMAN SPLIC	3.87
	452277	AL049013	Hs.28783	KIAA1223 protein	3.87
	438279	AA805166	Hs.165165	HIV-1 rev binding protein 2	3.86
20	458229	AI929602	Hs.177	phosphatidylinositol glycan, class H	3.86
	406414				3.86
	417193	AI922189	Hs.288390	hypothetical protein FLJ22795	3.85
	413174	AA723564	Hs.191343	ESTs	3.85
	433332	AI367347	Hs.127809	Homo sapiens clone TCCCTA00151 mRNA sequ	3.85
25	411089	AA456454	Hs.118637	cell division cycle 2-like 1 (PITSLRE pr	3.85
	412494	AL133900	Hs.792	ADP-ribosylation factor domain protein 1	3.84
	413530	AA130158	Hs.19977	ESTs, Moderately similar to ALU8_HUMAN A	3.84
	459592	AL037421	Hs.208746	ESTs, Moderately similar to pot. ORF I [	3.84
	418329	AW247430	Hs.84152	cystathionine-beta-synthase	3.83
30	451468	AW503398	Hs.210047	ESTs, Moderately similar to I38022 hypot	3.83
	434804	AA649530		gb:ns44f05.s1 NCI_CGAP_Alv1 Homo sapiens	3.83
	401819				3.82
	424179	F30712		Homo sapiens, clone IMAGE:4285740, mRNA	3.82
	424850	AA151057	Hs.153498	chromosome 18 open reading frame 1	3.82
35	426472	BE246138	Hs.30853	ESTs	3.82
	426625	T78300	Hs.171409	serologically defined colon cancer antig	3.82
	427585	D31152	Hs.179729	collagen, type X, alpha 1 (Schmid metaph	3.82
	427756	AI376540	Hs.15574	ESTs	3.82
	444701	AI916512	Hs.198394	ESTs	3.82
40	423052	M28214	Hs.123072	RAB3B, member RAS oncogene family	3.82
	429259	AA420450	Hs.292911	ESTs, Highly similar to S60712 band-6-pr	3.82
	416111	AA033813	Hs.79018	chromatin assembly factor 1, subunit A (	3.82
	433586	T85301		gb:yd78d06.s1 Soares fetal liver spleen	3.81
	438527	AI969251	Hs.143237	RAB7, member RAS oncogene family-like 1	3.81
45	410297	AA148710	Hs.159441	lumican	3.81
	429898	AW117322	Hs.42366	ESTs	3.81
	409079	W87707	Hs.82065	interleukin 6 signal transducer (gp130,	3.80
	419423	D26488	Hs.90315	KIAA0007 protein	3.80
	429643	AA455889	Hs.187548	FYVE-finger-containing Rab5 effector pro	3.80
50	431499	NM_001514	Hs.258561	general transcription factor IIB	3.80
	445060	AA830811	Hs.88803	ESTs	3.80
	449419	R34910	Hs.119172	ESTs	3.80
	450584	AA040403	Hs.60371	ESTs	3.80
	426137	AL040683	Hs.167031	DKFZP566D133 protein	3.79
55	420185	AL044056	Hs.158047	ESTs	3.79
	410076	T05387	Hs.7991	ESTs	3.78
	444078	BE246919	Hs.10290	U5 snRNP-specific 40 kDa protein (hPrp8-	3.78
	417318	AW953937	Hs.12891	ESTs	3.78
	414664	AA587775	Hs.66295	multi-PDZ-domain-containing protein	3.78
60	410275	U85658	Hs.61796	transcription factor AP-2 gamma (activat	3.77
	410503	AW975746	Hs.188662	KIAA1702 protein	3.77
	434170	AA626509	Hs.122329	ESTs	3.77
	421838	AW881089	Hs.108806	Homo sapiens mRNA; cDNA DKFZp566M0947 (f	3.77
	425268	AI807883	Hs.156932	Homo sapiens cDNA FLJ20653 fis, clone KA	3.76
65	431696	AA259068	Hs.267819	protein phosphatase 1, regulatory (inhib	3.76
	411990	AW963624	Hs.31707	ESTs, Weakly similar to YEW4_YEAST HYPOT	3.76
	430291	AV660345	Hs.238126	CGI-49 protein	3.76
	448779	BE042877	Hs.177135	ESTs	3.76
	452682	AA456193	Hs.155606	progesterone membrane binding protein	3.75

	452598	AI831594	Hs.68647	ESTs, Weakly similar to ALU7_HUMAN ALU S	3.75
	439498	AA908731	Hs.58297	CLLL8 protein	3.75
	440258	AI741633	Hs.125350	ESTs	3.74
	456848	AL121087	Hs.296406	KIAA0685 gene product	3.74
5	415082	AA160000	Hs.137396	ESTs, Weakly similar to JC5238 galactosy	3.74
	420653	AI224532	Hs.88550	ESTs	3.74
	431637	AI879330	Hs.265960	hypothetical protein FLJ10563	3.74
	440411	N30256	Hs.156971	hypothetical protein DKFZp434G1415	3.74
	405917				3.74
10	419440	AB020689	Hs.90419	KIAA0882 protein	3.74
	451230	BE546208	Hs.26090	hypothetical protein FLJ20272	3.73
	429597	NM_003816	Hs.2442	a disintegrin and metalloproteinase doma	3.73
	430144	AI732722	Hs.187694	ERGL protein; ERGIC-53-like protein	3.72
	438394	BE379623	Hs.27693	peptidylprolyl isomerase (cyclophilin)-l	3.72
15	440527	AV657117	Hs.184164	ESTs, Moderately similar to S65657 alpha	3.72
	449433	AI672096	Hs.9012	ESTs, Weakly similar to S26650 DNA-bindi	3.72
	456228	BE503227	Hs.134759	ESTs	3.72
	448663	BE614599	Hs.106823	hypothetical protein MGC14797	3.72
	415075	L27479	Hs.77889	Friedreich ataxia region gene X123	3.72
20	433544	AI793211	Hs.165372	ESTs, Moderately similar to ALU1_HUMAN A	3.71
	418293	AI224483	Hs.16063	hypothetical protein FLJ21877	3.71
	449897	AW819642	Hs.24135	transmembrane protein vezatin; hypotheti	3.71
	420297	AI626272	Hs.88323	ESTs, Weakly similar to ALU1_HUMAN ALU S	3.70
	423065	R96158	Hs.194606	Homo sapiens, clone MGC:5406, mRNA, comp	3.70
25	429340	N35938	Hs.199429	Homo sapiens mRNA; cDNA DKFZp434M2216 (f	3.70
	437777	AA768098	Hs.189079	ESTs	3.70
	440351	AF030933	Hs.7179	RAD1 (S. pombe) homolog	3.70
	443603	BE502601	Hs.134289	ESTs, Weakly similar to KIAA1063 protein	3.70
	446965	BE242873	Hs.16677	WD repeat domain 15	3.70
30	412350	AI659306	Hs.73826	protein tyrosine phosphatase, non-recept	3.70
	433852	AI376329	Hs.126629	ESTs	3.70
	433142	AL120697	Hs.110640	ESTs	3.69
	419994	AA282881	Hs.190057	ESTs	3.69
	412628	AI972402	Hs.173902	hypothetical protein MGC2648	3.69
35	431416	AA532718	Hs.178604	ESTs	3.69
	439444	AI277652	Hs.54578	ESTs, Weakly similar to I38022 hypotheti	3.68
	414709	AA704703	Hs.77031	Sp2 transcription factor	3.68
	447397	BE247676	Hs.18442	E-1 enzyme	3.68
	405718				3.68
40	425217	AU076696	Hs.155174	CDC5 (cell division cycle 5, S. pombe, h	3.68
	442242	AV647908	Hs.90424	Homo sapiens cDNA: FLJ23285 fis, clone H	3.68
	424690	BE538356	Hs.151777	eukaryotic translation initiation factor	3.68
	421734	AI318624	Hs.107444	Homo sapiens cDNA FLJ20562 fis, clone KA	3.67
	427221	L15409	Hs.174007	von Hippel-Lindau syndrome	3.67
45	439864	AI720078	Hs.291997	ESTs, Weakly similar to A47582 B-cell gr	3.66
	402408				3.66
	426327	W03242	Hs.44898	Homo sapiens clone TCCCTA00151 mRNA sequ	3.66
	427119	AW880562	Hs.114574	ESTs	3.66
	427356	AW023482	Hs.97849	ESTs	3.66
50	452946	X95425	Hs.31092	EphA5	3.66
	419078	M93119	Hs.89584	insulinoma-associated 1	3.66
	416295	AI064824	Hs.193385	ESTs	3.65
	427144	X95097	Hs.2126	vasoactive intestinal peptide receptor 2	3.65
	447500	AI381900	Hs.159212	ESTs	3.65
55	453127	AI696671	Hs.294110	ESTs	3.65
	423396	AI382555	Hs.127950	bromodomain-containing 1	3.65
	419346	AI830417		polybromo 1	3.64
	441540	C01367	Hs.127128	ESTs	3.64
	446501	AI302616	Hs.150819	ESTs	3.64
60	459527	AW977556	Hs.291735	ESTs, Weakly similar to I78885 serine/th	3.63
	446320	AF126245	Hs.14791	acyl-Coenzyme A dehydrogenase family, me	3.63
	435706	W31254	Hs.7045	GL004 protein	3.63
	400110				3.62
	410313	R10305	Hs.185683	ESTs	3.62
65	414713	BE465243	Hs.12664	ESTs	3.62
	436279	AW900372	Hs.180793	ESTs, Weakly similar to S65657 alpha-1C-	3.62
	439818	AL360137	Hs.19934	Homo sapiens mRNA full length insert cDN	3.62
	451797	AW663858	Hs.56120	small inducible cytokine subfamily E, me	3.62
	451294	AI457338	Hs.29894	ESTs	3.62

	434194	AF119847	Hs.283940	Homo sapiens PRO1550 mRNA, partial cds	3.62
	404939				3.62
	408101	AW968504	Hs.123073	CDC2-related protein kinase 7	3.62
	435846	AA700870	Hs.14304	ESTs	3.61
5	432833	N51075	Hs.47191	ESTs	3.61
	427276	AA400269	Hs.49598	ESTs	3.61
	433495	AW373784	Hs.71	alpha-2-glycoprotein 1, zinc	3.60
	403137				3.60
	404165				3.60
10	409571	AA504249	Hs.187585	ESTs	3.60
	410561	BE540255	Hs.6994	Homo sapiens cDNA: FLJ22044 fis, clone H	3.60
	412924	BE018422	Hs.75258	H2A histone family, member Y	3.60
	434228	Z42047	Hs.283978	Homo sapiens PRO2751 mRNA, complete cds	3.60
	436797	AA731491	Hs.178518	hypothetical protein MGC14879	3.60
15	437162	AW005505	Hs.5464	thyroid hormone receptor coactivating pr	3.60
	437444	H46008	Hs.31518	ESTs	3.60
	404210				3.59
	446157	BE270828	Hs.131740	Homo sapiens cDNA: FLJ22562 fis, clone H	3.59
	437587	AI591222	Hs.122421	Human DNA sequence from clone RP1-187J11	3.58
20	423147	AA987927	Hs.131740	Homo sapiens cDNA: FLJ22562 fis, clone H	3.57
	452226	AA024898	Hs.296002	ESTs	3.56
	443775	AF291664	Hs.204732	matrix metalloproteinase 26	3.56
	452501	AB037791	Hs.29716	hypothetical protein FLJ10980	3.56
	428647	AA830050	Hs.124344	ESTs	3.56
25	422443	NM_014707	Hs.116753	histone deacetylase 7B	3.55
	447966	AA340605	Hs.105887	ESTs, Weakly similar to Homolog of rat Z	3.55
	420892	AW975076	Hs.172589	nuclear phosphoprotein similar to S. cer	3.55
	420230	AL034344	Hs.298020	forkhead box C1	3.55
	418428	Y12490	Hs.85092	thyroid hormone receptor interactor 11	3.54
30	428949	AA442153	Hs.104744	hypothetical protein DKFZp434J0617	3.54
	444929	AI685841	Hs.161354	ESTs	3.54
	433339	AF019226	Hs.8036	glioblastoma overexpressed	3.54
	424369	R87622	Hs.26714	KIAA1831 protein	3.54
	433002	AF048730	Hs.279906	cyclin T1	3.53
35	435425	H16263	Hs.31416	ESTs	3.53
	415621	AI648602	Hs.131189	ESTs	3.53
	416974	AF010233	Hs.80667	RALBP1 associated Eps domain containing	3.53
	405793				3.52
	409770	AW499536		gb:UI-HF-BR0p-aji-c-12-0-UI.r1 NIH_MGC_5	3.52
40	425305	AA363025	Hs.155572	Human clone 23801 mRNA sequence	3.52
	428939	AW236550	Hs.131914	ESTs	3.52
	438388	AA806349	Hs.44698	ESTs	3.52
	443703	AV646177	Hs.213021	ESTs	3.52
	457940	AL360159	Hs.30445	Homo sapiens TRIpartite motif protein ps	3.52
45	402444				3.52
	409643	AW450866	Hs.257359	ESTs	3.51
	418250	U29926	Hs.83918	adenosine monophosphate deaminase (isofo	3.51
	432745	AI821926	Hs.269507	gb:nt78f05.x5 NCI_CGAP_Pr3 Homo sapiens	3.51
	414222	AL135173	Hs.878	sorbitol dehydrogenase	3.51
50	430061	AB037817	Hs.230188	KIAA1396 protein	3.51
	421491	H99999	Hs.42736	ESTs	3.50
	422384	AA224077	Hs.42438	Sm protein F	3.50
	434565	T52172		ESTs	3.50
	438379	N23018	Hs.171391	C-terminal binding protein 2	3.50
55	439741	BE379646	Hs.6904	Homo sapiens mRNA full length insert cDN	3.50
	447311	R37010	Hs.33417	Homo sapiens cDNA: FLJ22806 fis, clone K	3.50
	447805	AW627932	Hs.19614	gemin4	3.50
	454265	H03556	Hs.300949	ESTs, Weakly similar to thyroid hormone	3.50
	418838	AW385224	Hs.35198	ectonucleotide pyrophosphatase/phosphodi	3.50
60	448804	AW512213	Hs.42500	ADP-ribosylation factor-like 5	3.50
	409617	BE003760	Hs.55209	Homo sapiens mRNA; cDNA DKFZp434K0514 (f	3.49
	434075	AW003416	Hs.160604	ESTs	3.49
	444190	AI878918	Hs.10526	cysteine and glycine-rich protein 2	3.49
	435017	AA336522	Hs.12854	angiotensin II, type I receptor-associat	3.48
65	423445	NM_014324	Hs.128749	alpha-methylacyl-CoA racemase	3.48
	420271	AI954365	Hs.42892	ESTs	3.48
	443684	AI681307	Hs.166674	ESTs	3.48
	444168	AW379879		gb:RC1-HT0256-081199-011-f01 HT0256 Homo	3.48
	446074	AA079799	Hs.29263	hypothetical protein FLJ11896	3.48

	452582	AL137407	Hs.29911	Homo sapiens mRNA; cDNA DKFZp434M232 (fr	3.48
	431542	H63010	Hs.5740	ESTs	3.48
	432697	AW975050	Hs.293892	ESTs, Weakly similar to ALU4_HUMAN ALU S	3.48
	435572	AW975339	Hs.239828	ESTs, Weakly similar to GAG2_HUMAN RETRO	3.47
5	407192	AA609200		gb:af12e02.s1 Soares_testis_NHT Homo sap	3.47
	413435	X51405	Hs.75360	carboxypeptidase E	3.46
	447210	AF035269	Hs.17752	phosphatidylserine-specific phospholipas	3.46
	447958	AW796524	Hs.68644	Homo sapiens microsomal signal peptidase	3.46
	425312	AA354940	Hs.145958	ESTs	3.46
10	442007	AA301116	Hs.142838	nucleolar phosphoprotein Nopp34	3.46
	417455	AW007066	Hs.18949	ESTs, Weakly similar to CA2B_HUMAN COLLA	3.45
	426931	NM_003416	Hs.2076	zinc finger protein 7 (KOX 4, clone HF.1	3.45
	408739	W01556	Hs.238797	ESTs, Moderately similar to I38022 hypot	3.45
	436024	AI800041	Hs.190555	ESTs	3.45
15	408418	AW963897	Hs.44743	KIAA1435 protein	3.45
	409151	AA306105	Hs.50785	SEC22, vesicle trafficking protein (S. c	3.44
	418626	AW299508	Hs.135230	ESTs	3.44
	420560	AW207748	Hs.59115	ESTs	3.44
	420686	AI950339	Hs.40782	ESTs	3.44
20	428870	AA436831	Hs.36049	ESTs	3.44
	436754	AI061288	Hs.133437	ESTs	3.44
	437960	AI669586	Hs.222194	ESTs	3.44
	452300	AW628045	Hs.28896	Homo sapiens mRNA full length insert cDN	3.44
	421887	AW161450	Hs.109201	CGI-86 protein	3.44
25					

TABLE 5A shows the accession numbers for those primekeys lacking a unigeneID in Tables 5, 6, and 7. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Pkey: Unique Eos probeset Identifier number
CAT number: Gene cluster number
Accession: Genbank accession numbers

Pkey	CAT number	Accession
407596	1003489_1	R86913 R86901 H25352 R01370 H43764 AW044451 W21298
408432	1058667_1	AW195262 R27868 AW811262
409752	115301_1	AW963990 AA078196 AW749482 AA077468 BE151571 AA376917
409770	1154048_1	AW499536 AW499553 AW502138 AW499537 AW502136 AW501743
411440	124577_1	AW749402 AW749403 Z45743 R80376 AA093358
411479	1247077_1	AW848047 AW848202 AW848631 AW848142 AW848702 AW848121 AW848632 AW848140 AW848571 AW848009 AW848067 AW848069 AW848905 AW848214
411624	1252166_1	BE145964 BE146286 AW854564
412991	134248_1	AW949013 AA126111
414269	143133_1	AA298489 AA137165
415123	1523390_1	D60925 D60828 D80787
415715	1548818_1	F30364 F36559 T15435
416288	1585983_1	H51299 H44619 H46391 R86024 H51892 T72744
416289	1586037_1	W26333 R05358 H44682
417730	1695795_1	Z44761 R25801 R11926 R35604
418636	177402_1	AW749855 AA225995 AW750208 AW750206
419346	184129_1	AI830417 AA236612
419536	185688_1	AA603305 AA244095 AA244183
420111	190755_1	AA255652 AA280911 AW967920 AA262684
422219	213547_1	AW976073 AW976072 AA807550 AA306567
424179	236389_1	F30712 F35665 AW263888 AI904014 AI904018 AA336927 AA336502
424242	237181_1	AA337476 AW966227 AA450376 AW960222 AA381051
428002	265602_1	AA418703 AA418711 BE071915 BE071920 BE071912
429163	300543_1	AA884766 AW974271 AA592975 AA447312
432189	342819_1	AA527941 AI810608 AI620190 AA635266
432340	345248_1	AA534222 AA632632 T81234
432363	345469_1	AA534489 AW970240 AW970323
432966	356839_1	AA650114 AW974148 AA572946
433586	370470_1	T85301 AW517087 AA601054 BE073959
433641	37186_1	AF080229 AF080231 AF080230 AF080232 AF080233 AF080234 BE550633 AI636743 AW614951 BE467547 AI680833 AI633818 N29986 U87592 U87593 U87590 U87591 S46404 U87587 AA463992 AW206802 AI970376 AI583718 AI672574 N25695 AW665466 AI818326 AA126128 AI480345 AW013827 AA248638 AI214968 AA204735 AA207155 AA206262 AA204833 AW003247 AW496808 AI080480 AI631703 AI651023 AI867418 AW818140 AA502500 AI206199 AI671282 AI352545 BE501030 AI652535 BE465762 AA206331 AW451866 AA471088 AA206342 AA204834 AA206100 AW021661 AA332922 N66048 AA703396 H92278 AW139734 H92683 U87589 U87595 H69001 U87594 BE466420 AI624817 BE466611 AI206344 AA574397 AA348354 AI493192
433687	373061_1	AA743991 AA604852 AW272737
433891	376239_1	AA613792 AW182329 T05304 AW858385
434415	385931_1	BE177494 AW276909 AA632849
434565	38898_1	T52172 AF147324 T52248
434804	393481_1	AA649530 AA659316 H64973
437113	433234_1	AA744693 AW750059
444168	593829_1	AW379879 AI126285 H12014
448212	755099_1	AI475858 AW969013
448310	757918_1	AI480316 AW847535
451746	883303_1	M86178 AI813822 D56993

452560	922216_1	BE077084 AW139963 AW863127 AW806209 AW806204 AW806205 AW806206 AW806211 AW806212 AW806207 AW806208 AW806210 AI907497
452712	928309_1	AW838616 AW838660 BE144343 AI914520 AW888910 BE184854 BE184784
453773	980699_1	AL133761 AL133767
455276	1272541_1	BE176479 BE176678 BE176357 BE176550 AW886079 BE176676 BE176615 BE176555 BE176489 BE176610 BE176362
455309	1278153_1	AW894017 AW893956 AW894032

TABLE 5B shows the genomic positioning for those primekeys lacking unigene ID's and accession numbers in Tables 5, 6, and 7. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

Pkey:	Unique number corresponding to an Eos probeset
Ref:	Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-495.
Strand:	Indicates DNA strand from which exons were predicted.
Nt_position:	Indicates nucleotide positions of predicted exons.

Pkey	Ref	Strand	Nt_position
401045	8117619	Plus	90044-90184,91111-91345
401424	8176894	Plus	24223-24428
401451	6634068	Minus	119926-121272
401714	6715702	Plus	96484-96681
401747	9789672	Minus	118596-118816,119119-119244,119609-119761,120422-120990,130161-130381,130468-130593,131097-131258,131866-131932,132451-132575,133580-134011
401785	7249190	Minus	165776-165996,166189-166314,166408-166569,167112-167268,167387-167469,168634-168942
401819	7467933	Minus	28217-28486
402408	9796239	Minus	110326-110491
402444	9796614	Plus	28391-28517
402791	6137008	Minus	51036-51207
403047	3540153	Minus	59793-59968
403137	9211494	Minus	92349-92572,92958-93084,93579-93712,93949-94072,94591-94748,95214-95337
403721	7528046	Minus	156647-157366
403764	7717105	Minus	118692-118853
403797	8099896	Minus	123065-125008
404165	9926489	Minus	69025-69128
404210	5006246	Plus	169926-170121
404253	9367202	Minus	55675-56055
404561	9795980	Minus	69039-70100
404571	7249169	Minus	112450-112648
404721	9856648	Minus	173763-174294
404915	7341766	Minus	100915-101087
404939	6862697	Plus	175318-175476
405403	6850244	Minus	37491-37670,40951-41031
405685	4508129	Minus	37956-38097
405718	9795467	Plus	113080-113266
405793	1405887	Minus	89197-89453
405876	6758747	Plus	39694-40031
405917	7712162	Minus	106829-107213
406414	9256407	Plus	49593-49850
406554	7711566	Plus	106956-107121

TABLE 6:286 GENES ENCODING EXTRACELLULAR OR CELL SURFACE PROTEINS UP-REGULATED IN PROSTATE CANCER COMPARED TO NORMAL ADULT TISSUES

Table 6 shows 286 genes up-regulated in prostate cancer compared to normal adult tissues that are likely to be extracellular or cell-surface proteins. These were selected as for Table 5 and the predicted protein contained a structural domain that is indicative of extracellular localization (e.g. egf, 7tm domains).

Pkey:	Unique Eos probeset identifier number
ExAccn:	Exemplar Accession number, Genbank accession number
UnigeneID:	Unigene number
Unigene Title:	Unigene gene title
R1:	Ratio of tumor to normal tissue

Pkey	ExAccn	UnigeneID	Uningene Title	R1
409361	NM_005982	Hs.54416	sine oculis homeobox (Drosophila) homolo	48.28
409731	AA125985	Hs.56145	thymosin, beta, identified in neuroblast	45.24
400298	AA032279	Hs.61635	six transmembrane epithelial antigen of	43.48
420154	AI093155	Hs.95420	JM27 protein	41.12
426747	AA535210	Hs.171995	kallikrein 3, (prostate specific antigen	31.80
400299	X07730	Hs.171995	kallikrein 3, (prostate specific antigen	24.91
425075	AA506324	Hs.1852	acid phosphatase, prostate	24.23
424846	AU077324	Hs.1832	neuropeptide Y	23.57
405685				20.90
420757	X78592	Hs.99915	androgen receptor (dihydrotestosterone r	19.72
418994	AA296520	Hs.89546	selectin E (endothelial adhesion molecul	19.56
452792	AB037765	Hs.30652	KIAA1344 protein	17.39
445472	AB006631	Hs.12784	Homo sapiens mRNA for KIAA0293 gene, par	17.00
414565	AA502972	Hs.183390	hypothetical protein FLJ13590	16.82
431716	D89053	Hs.268012	fatty-acid-Coenzyme A ligase, long-chain	16.60
408430	S79876	Hs.44926	dipeptidylpeptidase IV (CD26, adenosine	16.28
408000	L11690	Hs.620	bullous pemphigoid antigen 1 (230/240kD)	15.54
430226	BE245562	Hs.2551	adrenergic, beta-2-, receptor, surface	15.40
444484	AK002126	Hs.11260	hypothetical protein FLJ11264	14.76
418601	AA279490	Hs.86368	calmegin	14.56
448999	AF179274	Hs.22791	transmembrane protein with EGF-like and	14.55
416182	NM_004354	Hs.79069	cyclin G2	12.94
420544	AA677577	Hs.98732	Homo sapiens Chromosome 16 BAC clone CIT	12.79
445413	AA151342	Hs.12677	CGI-147 protein	12.64
453930	AA419466	Hs.36727	hypothetical protein FLJ10903	12.22
440286	U29589	Hs.7138	cholinergic receptor, muscarinic 3	12.04
452784	BE463857	Hs.151258	hypothetical protein FLJ21062	11.86
450203	AF097994	Hs.301528	L-kynurenine/alpha-aminoadipate aminotra	11.68
448045	AJ297436	Hs.20166	prostate stem cell antigen	11.51
449650	AF055575	Hs.23838	calcium channel, voltage-dependent, L ty	11.18
420381	D50640	Hs.337616	phosphodiesterase 3B, cGMP-inhibited	11.10
425665	AK001050	Hs.159066	hypothetical protein FLJ10188	11.08
425710	AF030880	Hs.159275	solute carrier family, member 4	11.08
428728	NM_016625	Hs.191381	hypothetical protein	11.04
407021	U52077		gb:Human mariner1 transposase gene, comp	11.02
410733	D84284	Hs.66052	CD38 antigen (p45)	11.02
452340	NM_002202	Hs.505	ISL1 transcription factor, LIM/homeodoma	10.85
428819	AL135623	Hs.193914	KIAA0575 gene product	10.48
421991	NM_014918	Hs.110488	KIAA0990 protein	10.04
431217	NM_013427	Hs.250830	Rho GTPase activating protein 6	9.75
421470	R27496	Hs.1378	annexin A3	9.64
409262	AK000631	Hs.52256	hypothetical protein FLJ20624	9.45
435980	AF274571	Hs.129142	deoxyribonuclease II beta	9.24
421246	AW582962	Hs.102897	CGI-47 protein	9.20
410001	AB041036	Hs.57771	kallikrein 11	9.03
441791	AW372449	Hs.175982	hypothetical protein FLJ21159	9.02

	404571				8.66
	456497	AW967956	Hs.123648	ESTs, Weakly similar to AF108460 1 ubinu	8.56
	419968	X04430	Hs.93913	interleukin 6 (interferon, beta 2)	8.36
	433172	AB037841	Hs.102652	hypothetical protein ASH1	8.30
5	422631	BE218919	Hs.118793	hypothetical protein FLJ10688	8.27
	427674	NM_003528	Hs.2178	H2B histone family, member Q	8.20
	404915				8.08
	452259	AA317439	Hs.28707	signal sequence receptor, gamma (transio	8.06
	452891	N75582	Hs.212875	ESTs, Weakly similar to DYH9_HUMAN CILIA	8.02
10	439731	AI953135	Hs.45140	hypothetical protein FLJ14084	7.98
	419839	U24577	Hs.93304	phospholipase A2, group VII (platelet-ac	7.68
	420120	AL049610	Hs.95243	transcription elongation factor A (SII)-	7.64
	424099	AF071202	Hs.139336	ATP-binding cassette, sub-family C (CFTR	7.64
	448706	AW291095	Hs.21814	interleukin 20 receptor, alpha	7.52
15	410227	AB009284	Hs.61152	exostoses (multiple)-like 2	7.49
	425211	M18667	Hs.1867	progastricsin (pepsinogen C)	7.35
	441736	AW292779	Hs.169799	ESTs	7.28
	419991	AJ000098	Hs.94210	eyes absent (Drosophila) homolog 1	7.20
	425018	BE245277	Hs.154196	E4F transcription factor 1	7.20
20	424560	AA158727	Hs.150555	protein predicted by clone 23733	7.18
	409110	AA191493	Hs.48778	niban protein	7.10
	421566	NM_000399	Hs.1395	early growth response 2 (Krox-20 (Drosop	7.04
	431725	X65724	Hs.2839	Norrie disease (pseudoglioma)	6.98
	425782	U66468	Hs.159525	cell growth regulatory with EF-hand doma	6.85
25	427408	AA583206	Hs.2156	RAR-related orphan receptor A	6.79
	435604	AA625279	Hs.26892	uncharacterized bone marrow protein BM04	6.73
	415874	AF091622	Hs.78893	KIAA0244 protein	6.54
	401451				6.52
	431778	AL080276	Hs.268562	regulator of G-protein signalling 17	6.51
30	409089	NM_014781	Hs.50421	KIAA0203 gene product	6.50
	431992	NM_002742	Hs.2891	protein kinase C, mu	6.49
	404253				6.42
	421552	AF026692	Hs.105700	secreted frizzled-related protein 4	6.41
	416806	NM_000288	Hs.79993	peroxisomal biogenesis factor 7	6.38
35	431958	X63629	Hs.2877	cadherin 3, type 1, P-cadherin (placenta	6.30
	439366	AF100143	Hs.6540	fibroblast growth factor 13	6.30
	416836	D54745	Hs.80247	cholecystokinin	6.30
	433383	AF034837	Hs.192731	double-stranded RNA specific adenosine d	6.29
	450728	AW162923	Hs.25363	presenilin 2 (Alzheimer disease 4)	6.25
40	413384	NM_000401	Hs.75334	exostoses (multiple) 2	6.22
	423349	AF010258	Hs.127428	homeo box A9	6.20
	424800	AL035588	Hs.153203	MyoD family inhibitor	6.18
	425451	AF242769	Hs.157461	mesenchymal stem cell protein DSC54	6.14
	447359	NM_012093	Hs.18268	adenylate kinase 5	6.00
45	410889	X91662	Hs.66744	twist (Drosophila) homolog (acrocephalos	5.97
	408829	NM_006042	Hs.48384	heparan sulfate (glucosamine) 3-O-sulfot	5.94
	453911	AW503857	Hs.4007	Sarcolemmal-associated protein	5.94
	408875	NM_015434	Hs.48604	DKFZP434B168 protein	5.92
	450480	X82125	Hs.25040	zinc finger protein 239	5.90
50	451684	AF216751	Hs.26813	CDA14	5.88
	400301	X03635	Hs.1657	estrogen receptor 1	5.78
	415077	L41607	Hs.934	glucosaminyl (N-acetyl) transferase 2, I	5.74
	418852	BE537037	Hs.273294	hypothetical protein FLJ20069	5.72
	446867	AB007891	Hs.16349	KIAA0431 protein	5.72
55	410232	AW372451	Hs.61184	CGI-79 protein	5.70
	422762	AL031320	Hs.119976	Human DNA sequence from clone RP1-20N2 o	5.70
	450616	AL133067	Hs.302689	hypothetical protein	5.70
	408621	AI970672	Hs.46638	chromosome 11 open reading frame 8	5.65
	439671	AW162840	Hs.6641	kinesin family member 5C	5.64
60	410196	AI936442	Hs.59838	hypothetical protein FLJ10808	5.60
	429170	NM_001394	Hs.2359	dual specificity phosphatase 4	5.60
	440738	AI004650	Hs.225674	WD repeat domain 9	5.60
	414342	AA742181	Hs.75912	KIAA0257 protein	5.59
	422634	NM_016010	Hs.118821	CGI-62 protein	5.56
65	400268				5.55
	439569	AW602166	Hs.222399	CEGP1 protein	5.51
	452823	AB012124	Hs.30696	transcription factor-like 5 (basic helix	5.48
	431938	AA938471	Hs.54431	specific granule protein (28 kDa); cyste	5.44
	427638	AA406411	Hs.208341	ESTs, Weakly similar to KIAA0989 protein	5.42

421264	AL039123	Hs.103042	microtubule-associated protein 1B	5.38
421685	AF189723	Hs.106778	ATPase, Ca++ transporting, type 2C, memb	5.37
421987	AI133161	Hs.286131	CGI-101 protein	5.36
422806	BE314767	Hs.1581	glutathione S-transferase theta 2	5.34
432281	AK001239	Hs.274263	hypothetical protein FLJ10377	5.32
451982	F13036	Hs.27373	Homo sapiens mRNA; cDNA DKFZp564O1763 (f	5.32
444042	NM_004915	Hs.10237	ATP-binding cassette, sub-family G (WHIT	5.31
447752	M73700	Hs.105938	lactotransferrin	5.29
451418	BE387790	Hs.26369	hypothetical protein FLJ20287	5.22
428593	AW207440	Hs.185973	degenerative spermatocyte (homolog Droso	5.21
447541	AK000288	Hs.18800	hypothetical protein FLJ20281	5.18
459294	AW977286	Hs.17428	RBP1-like protein	5.16
424692	AA429834	Hs.151791	KIAA0092 gene product	5.15
416434	AW163045	Hs.79334	nuclear factor, interleukin 3 regulated	5.11
410268	AA316181	Hs.61635	six transmembrane epithelial antigen of	5.10
417517	AF001176	Hs.82238	POP4 (processing of precursor , S. cerev	5.10
453616	NM_003462	Hs.33846	dynein, axonemal, light intermediate pol	5.10
427958	AA418000	Hs.98280	potassium intermediate/small conductance	5.09
407945	X69208	Hs.606	ATPase, Cu++ transporting, alpha polypep	5.08
418576	AW968159	Hs.289104	Alu-binding protein with zinc finger dom	5.05
413328	Y15723	Hs.75295	guanylate cyclase 1, soluble, alpha 3	5.04
432729	AK000292	Hs.278732	hypothetical protein FLJ20285	5.04
426342	AF093419	Hs.169378	multiple PDZ domain protein	5.02
429782	NM_005754	Hs.220689	Ras-GTPase-activating protein SH3-domain	5.02
436209	AW850417	Hs.254020	ESTs, Moderately similar to unnamed prot	5.02
430599	NM_004855	Hs.247118	phosphatidylinositol glycan, class B	5.00
451386	AB029006	Hs.26334	spastic paraplegia 4 (autosomal dominant	5.00
457211	AW972565	Hs.32399	ESTs, Weakly similar to S51797 vasodilat	4.97
425851	NM_001490	Hs.159642	glucosaminyl (N-acetyl) transferase 1, c	4.97
421689	N87820	Hs.106826	KIAA1696 protein	4.93
416533	BE244053	Hs.79362	retinoblastoma-like 2 (p130)	4.92
432653	N62096	Hs.293185	ESTs, Weakly similar to JC7328 amino aci	4.91
403047				4.91
431117	AF003522	Hs.250500	delta (Drosophila)-like 1	4.90
427617	D42063	Hs.199179	RAN binding protein 2	4.88
428804	AK000713	Hs.193736	hypothetical protein FLJ20706	4.88
449071	NM_005872	Hs.22960	breast carcinoma amplified sequence 2	4.86
407596	R86913		gb:yq30f05.r1 Soares fetal liver spleen	4.84
456516	BE172704	Hs.222746	KIAA1610 protein	4.84
458339	AW976853	Hs.172843	ESTs	4.83
422083	NM_001141	Hs.111256	arachidonate 15-lipoxygenase, second typ	4.82
449535	W15267	Hs.23672	low density lipoprotein receptor-related	4.82
422048	NM_012445	Hs.288126	spondin 2, extracellular matrix protein	4.82
424602	AK002055	Hs.151046	hypothetical protein FLJ11193	4.78
410765	AI694972	Hs.66180	nucleosome assembly protein 1-like 2	4.77
419879	Z17805	Hs.93564	Homer, neuronal immediate early gene, 2	4.74
450649	NM_001429	Hs.25272	E1A binding protein p300	4.74
411624	BE145964	Hs.103283	KIAA0594 protein	4.72
404721				4.70
426261	AW242243	Hs.168670	peroxisomal farnesylated protein	4.70
416276	U41060	Hs.79136	LIV-1 protein, estrogen regulated	4.64
408374	AW025430	Hs.155591	forkhead box F1	4.64
451900	AB023199	Hs.27207	KIAA0982 protein	4.63
421437	AW821252	Hs.104336	hypothetical protein	4.63
434629	AA789081	Hs.4029	glioma-amplified sequence-41	4.60
403764				4.58
421247	BE391727	Hs.102910	general transcription factor IIH, polype	4.53
403721				4.50
453070	AK001465	Hs.31575	SEC63, endoplasmic reticulum translocon	4.49
417412	X16896	Hs.82112	interleukin 1 receptor, type I	4.48
439735	AI635386	Hs.142846	hypothetical protein	4.48
430261	AA305127	Hs.237225	hypothetical protein HT023	4.46
430598	AK001764	Hs.247112	hypothetical protein FLJ10902	4.44
400303	AA242758	Hs.79136	LIV-1 protein, estrogen regulated	4.42
438209	AL120659	Hs.6111	aryl-hydrocarbon receptor nuclear transl	4.42
417421	AL138201	Hs.82120	nuclear receptor subfamily 4, group A, m	4.40
447270	AC002551	Hs.331	general transcription factor IIIC, polyp	4.38
434423	NM_006769	Hs.3844	LIM domain only 4	4.35
404561				4.32

	422969	AA782536	Hs.122647	N-myristoyltransferase 2	4.32
	423685	BE350494	Hs.49753	uveal autoantigen with coiled coil domai	4.32
	425071	NM_013989	Hs.154424	deiodinase, iodothyronine, type II	4.32
	431583	AL042613	Hs.262476	S-adenosylmethionine decarboxylase 1	4.31
5	442818	AK001741	Hs.8739	hypothetical protein FLJ10879	4.30
	423740	Y07701	Hs.293007	aminopeptidase puromycin sensitive	4.24
	424701	NM_005923	Hs.151988	mitogen-activated protein kinase kinase	4.21
	424085	NM_002914	Hs.139226	replication factor C (activator 1) 2 (40	4.20
	410294	AB014515	Hs.323712	KIAA0615 gene product	4.18
10	447124	AW976438	Hs.17428	RBP1-like protein	4.18
	438018	AK001160	Hs.5999	hypothetical protein FLJ10298	4.16
	443857	AI089292	Hs.287621	hypothetical protein FLJ14069	4.15
	446711	AF169692	Hs.12450	protocadherin 9	4.15
	405403				4.14
15	448148	NM_016578	Hs.20509	HBV pX associated protein-8	4.13
	417531	NM_003157	Hs.1087	serine/threonine kinase 2	4.12
	433345	AI681545	Hs.152982	hypothetical protein FLJ13117	4.10
	432712	AB016247	Hs.288031	sterol-C5-desaturase (fungal ERG3, delta	4.09
	435114	AA775483	Hs.288936	mitochondrial ribosomal protein L9	4.08
20	445459	AI478629	Hs.158465	likely ortholog of mouse putative IKK re	4.08
	402791				4.04
	438660	U95740	Hs.6349	Homo sapiens, clone IMAGE:3010666, mRNA,	4.04
	447568	AF155655	Hs.18885	CGI-116 protein	4.04
	452211	AI985513	Hs.233420	ESTs	4.02
25	443292	AK000213	Hs.9196	hypothetical protein	4.01
	420911	U77413	Hs.100293	O-linked N-acetylglucosamine (GlcNAc) tr	4.00
	428738	NM_000380	Hs.192803	xeroderma pigmentosum, complementation g	3.95
	430456	AA314998	Hs.241503	hypothetical protein	3.95
	437531	AI400752	Hs.112259	T cell receptor gamma locus	3.93
30	428695	AI355647	Hs.189999	purinergic receptor (family A group 5)	3.91
	410011	AB020641	Hs.57856	PFTAIRE protein kinase 1	3.91
	446494	AA463276	Hs.288906	WW Domain-Containing Gene	3.91
	409928	AL137163	Hs.57549	hypothetical protein dJ473B4	3.90
	411598	BE336654	Hs.70937	H3 histone family, member A	3.90
35	425707	AF115402	Hs.11713	E74-like factor 5 (ets domain transcript	3.90
	451806	NM_003729	Hs.27076	RNA 3'-terminal phosphate cyclase	3.89
	401045				3.89
	437372	AA323968	Hs.283631	hypothetical protein DKFZp547G183	3.89
	417067	AJ001417	Hs.81086	solute carrier family 22 (extraneuronal	3.88
40	410467	AF102546	Hs.63931	dachshund (Drosophila) homolog	3.88
	431930	AB035301	Hs.272211	cadherin 7, type 2	3.88
	453047	AW023798	Hs.286025	ESTs	3.88
	401785				3.88
	458229	AI929602	Hs.177	phosphatidylinositol glycan, class H	3.86
45	406414				3.86
	412494	AL133900	Hs.792	ADP-ribosylation factor domain protein 1	3.84
	418329	AW247430	Hs.84152	cystathionine-beta-synthase	3.83
	424850	AA151057	Hs.153498	chromosome 18 open reading frame 1	3.82
	427585	D31152	Hs.179729	collagen, type X, alpha 1 (Schmid metaph	3.82
50	423052	M28214	Hs.123072	RAB3B, member RAS oncogene family	3.82
	416111	AA033813	Hs.79018	chromatin assembly factor 1, subunit A (	3.82
	419423	D26488	Hs.90315	KIAA0007 protein	3.80
	429643	AA455889	Hs.167279	FYVE-finger-containing Rab5 effector pro	3.80
	431499	NM_001514	Hs.258561	general transcription factor IIB	3.80
55	444078	BE246919	Hs.10290	U5 snRNP-specific 40 kDa protein (hPrp8-	3.78
	430291	AV660345	Hs.238126	CGI-49 protein	3.76
	431637	AI879330	Hs.265960	hypothetical protein FLJ10563	3.74
	440411	N30256	Hs.151093	hypothetical protein DKFZp434G1415	3.74
	405917				3.74
60	451230	BE546208	Hs.26090	hypothetical protein FLJ20272	3.73
	429597	NM_003816	Hs.2442	a disintegrin and metalloproteinase doma	3.73
	415075	L27479	Hs.77889	Friedreich ataxia region gene X123	3.72
	440351	AF030933	Hs.7179	RAD1 (S. pombe) homolog	3.70
	443603	BE502601	Hs.134289	ESTs, Weakly similar to KIAA1063 protein	3.70
65	446965	BE242873	Hs.16677	WD repeat domain 15	3.70
	412350	AI659306	Hs.73826	protein tyrosine phosphatase, non-recept	3.70
	433852	AI378329	Hs.126629	ESTs	3.70
	447397	BE247676	Hs.18442	E-1 enzyme	3.68
	405718				3.68

	425217	AU076696	Hs.155174	CDC5 (cell division cycle 5, S. pombe, h	3.68
	421734	AI318624	Hs.107444	Homo sapiens cDNA FLJ20562 fis, clone KA	3.67
	427221	L15409	Hs.174007	von Hippel-Lindau syndrome	3.67
	402408				3.66
5	452946	X95425	Hs.31092	EphA5	3.66
	419078	M93119	Hs.89584	insulinoma-associated 1	3.66
	427144	X95097	Hs.2126	vasoactive intestinal peptide receptor 2	3.65
	423396	AI382555	Hs.127950	bromodomain-containing 1	3.65
	446320	AF126245	Hs.14791	acyl-Coenzyme A dehydrogenase family, me	3.63
10	404939				3.62
	403137				3.60
	437162	AW005505	Hs.5464	thyroid hormone receptor coactivating pr	3.60
	404210				3.59
	443775	AF291664	Hs.204732	matrix metalloproteinase 26	3.56
15	452501	AB037791	Hs.29716	hypothetical protein FLJ10980	3.56
	422443	NM_014707	Hs.116753	histone deacetylase 7B	3.55
	420230	AL034344	Hs.284186	forkhead box C1	3.55
	418428	Y12490	Hs.85092	thyroid hormone receptor interactor 11	3.54
	433002	AF048730	Hs.279906	cyclin T1	3.53
20	405793				3.52
	457940	AL360159	Hs.306517	Homo sapiens TRIpartite motif protein ps	3.52
	402444				3.52
	418250	U29926	Hs.83918	adenosine monophosphate deaminase (isofo	3.51
	414222	AL135173	Hs.878	sorbitol dehydrogenase	3.51
25	422384	AA224077	Hs.42438	Sm protein F	3.50
	447805	AW627932	Hs.19614	gemin4	3.50
	454265	H03556	Hs.300949	ESTs, Weakly similar to thyroid hormone	3.50
	423445	NM_014324	Hs.128749	alpha-methylacyl-CoA racemase	3.48
	413435	X51405	Hs.75360	carboxypeptidase E	3.46
30	447210	AF035269	Hs.17752	phosphatidylserine-specific phospholipas	3.46
	426931	NM_003416	Hs.2076	zinc finger protein 7 (KOX 4, clone HF.1	3.45
	408418	AW963897	Hs.44743	KIAA1435 protein	3.45
	421887	AW161450	Hs.109201	CGI-86 protein	3.44

Table 7: 42 GENES ENCODING SMALL MOLECULE TARGETS UP-REGULATED IN PROSTATE CANCER COMPARED TO NORMAL ADULT TISSUES

Table 7 shows 42 genes up-regulated in prostate cancer compared to normal adult tissues that are likely to be small molecule targets. These were selected as for Table 5 and the predicted protein contained a structural domain that is indicative of a drugable structure (e.g. protease, kinase, phosphatase, receptor). The functional domain is indicated for each gene.

Pkey:	Unique Eos probeset identifier number
ExAccn:	Exemplar Accession number, Genbank accession number
UnigeneID:	Unigene number
Unigene Title:	Unigene gene title
PSDomain:	Protein Structural Domain
R1:	Ratio of tumor vs. normal tissue

Pkey	ExAccn	UnigeneID	Unigene Title	PSDomain	R1	
426747	AA535210	Hs.171995	kallikrein 3, (prostate specific antigen	trypsin	31.80	
400299	X07730	Hs.171995	kallikrein 3, (prostate specific antigen	trypsin	24.91	
420757	X78592	Hs.99915	androgen receptor (dihydrotestosterone r	Androgen_recep,hormone_rec,zf-C4	19.72	
408430	S79876	Hs.44926	dipeptidylpeptidase IV (CD26, adenosine	DPPIV_N_term,Peptidase_S9	16.28	
430226	BE245562	Hs.2551	adrenergic, beta-2-, receptor, surface	7tm_1	15.40	
411096	U80034	Hs.68583	mitochondrial intermediate peptidase	Peptidase_M3	14.81	
440286	U29589	Hs.7138	cholinergic receptor, muscarinic 3	7tm_1	12.04	
420381	D50640	Hs.337616	phosphodiesterase 3B, cGMP-inhibited	PDEase	11.10	
407021	U52077		gb:Human mariner1 transposase gene, comp	SET,Transposase_1	11.02	
401424				arginase	9.58	
410001	AB041036	Hs.57771	kallikrein 11	trypsin	9.03	
428330	L22524	Hs.2256	matrix metalloproteinase 7 (matrilysin,	Peptidase_M10	8.76	
424099	AF071202	Hs.139336	ATP-binding cassette, sub-family C (CFTR	ABC_tran,ABC_membrane	7.64	
419991	AJ000098	Hs.94210	eyes absent (Drosophila) homolog 1	Hydrolase	7.20	
431992	NM_002742	Hs.2891	protein kinase C, mu	pkinase,DAG_PE-bind,PH	6.49	
447359	NM_012093	Hs.18268	adenylate kinase 5	adenylatekinase	6.00	
400301	X03635	Hs.1657	estrogen receptor 1	Oest_recep,zf-C4,hormone_rec	5.78	
421685	AF189723	Hs.106778	ATPase, Ca++ transporting, type 2C, memb	E1-E2_ATPase,Hydrolase	5.37	
444042	NM_004915	Hs.10237	ATP-binding cassette, sub-family G (WHIT	ABC_tran	5.31	
447752	M73700	Hs.105938	lactotransferrin	transferrin,7tm_1	5.29	
407945	X69208	Hs.606	ATPase, Cu++ transporting, alpha polypep	E1-E2_ATPase,Hydrolase,HMA	5.08	
403047				trypsin	4.91	
427617	D42063	Hs.199179	RAN binding protein 2	Ran_BP1,zf-RanBP,TPR,pro_isomerase		4.88
422083	NM_001141	Hs.111256	arachidonate 15-lipoxygenase, second typ	lipoxygenase,PLAT	4.82	
449535	W15267	Hs.23672	low density lipoprotein receptor-related	ldl_recept_b,ldl_recept_a,EGF	4.82	
425071	NM_013989	Hs.154424	deiodinase, iodothyronine, type II	T4_deiodinase	4.32	
423740	Y07701	Hs.293007	aminopeptidase puromycin sensitive	Peptidase_M1	4.24	
424701	NM_005923	Hs.151988	mitogen-activated protein kinase kinase	pkinase	4.21	
424085	NM_002914	Hs.139226	replication factor C (activator 1) 2 (40	AAA,Viral_helicase1	4.20	
417531	NM_003157	Hs.1087	serine/threonine kinase 2	pkinase	4.12	
428695	AI355647	Hs.189999	purinergic receptor (family A group 5)	7tm_1	3.91	
410011	AB020641	Hs.57856	PFTAIRE protein kinase 1	pkinase	3.91	
424850	AA151057	Hs.153498	chromosome 18 open reading frame 1	ldl_recept_a	3.82	
412350	AI659306	Hs.73826	protein tyrosine phosphatase, non-recept	Y_phosphatase,Band_41,PDZ	3.70	
447397	BE247676	Hs.18442	E-1 enzyme	Hydrolase	3.68	
452946	X95425	Hs.31092	EphA5	EPH_lbd,fn3,pkinase,SAM	3.66	
427144	X95097	Hs.2126	vasoactive intestinal peptide receptor 2	7tm_2	3.65	
443775	AF291664	Hs.204732	matrix metalloproteinase 26	Peptidase_M10	3.56	
457940	AL360159	Hs.306517	Homo sapiens TRIpartite motif protein ps	SPRY,7tm_1	3.52	
418250	U29926	Hs.83918	adenosine monophosphate deaminase (isofo	A_deaminase	3.51	
413435	X51405	Hs.75360	carboxypeptidase E	Zn_carbOpept	3.46	
447210	AF035269	Hs.17752	phosphatidylserine-specific phospholipas	lipase	3.46	

TABLE 8: 136 GENES SIGNIFICANTLY DOWN-REGULATED IN PROSTATE CANCER COMPARED TO NORMAL PROSTATE

Table 8 shows 136 genes significantly down-regulated in prostate cancer compared to normal prostate . These were selected from 59680 probesets on the Affymetrix/Eos Hu03 GeneChip array such that the ratio of "average" normal prostate to "average" prostate cancer tissues was greater than or equal to 2. The "average" normal prostate level was set to the mean amongst 4 normal prostate tissues. The "average" prostate cancer level was set to the 85th percentile amongst 73 tumor samples. In order to remove gene-specific background levels of non-specific hybridization, the 10th percentile value amongst all the tissues was subtracted from both the numerator and the denominator before the ratio was evaluated.

Pkey:	Unique Eos probeset identifier number
ExAccn:	Exemplar Accession number, Genbank accession number
UnigeneID:	Unigene number
Unigene Title:	Unigene gene title
R1:	Ratio of normal prostate to prostate cancer

Pkey	ExAccn	UnigeneID	Unigene Title	R1
425932	M81650	Hs.1968	semenogelin I	57.69
425545	N98529	Hs.158295	Human mRNA for myosin light chain 3 (MLC	19.70
426752	X69490	Hs.172004	titin	15.25
442082	R41823	Hs.7413	ESTs; calsyntenin-2	10.05
407245	X90568	Hs.172004	titin	9.38
422711	D60641	Hs.21739	Homo sapiens mRNA; cDNA DKFZp586I1518 (f	9.05
420813	X51501	Hs.99949	prolactin-induced protein	8.18
411987	AA375975	Hs.183380	"ESTs, Moderately similar to ALU7_HUMAN	7.45
404567				5.62
416030	H15261	Hs.21948	ESTs	5.51
444892	AI620617	Hs.148565	ESTs	5.27
444573	AW043590	Hs.225023	ESTs	5.20
428068	AW016437	Hs.233462	ESTs	5.08
437440	AA846804	Hs.123694	ESTs	4.95
404113				4.75
452279	AA286844	Hs.61260	hypothetical protein FLJ13164	4.75
421058	AW297967	Hs.188181	ESTs	4.63
445592	AV654382	Hs.17947	"ESTs, Weakly similar to K02F3.10 [C.ele	4.53
405163				4.49
405227				4.45
454059	NM_003154	Hs.37048	statherin	4.45
450152	AI138635	Hs.22968	ESTs	4.40
407013	U35637		"gb:Human nebulin mRNA, partial cds"	4.03
403612				4.02
440089	AA864468	Hs.135646	ESTs	4.00
408988	AL119844	Hs.49476	Homo sapiens clone TUA8 Crl-du-chat regi	3.98
436726	AA324975	Hs.128993	"ESTs, Weakly similar to KIAA0465 protei	3.95
459367	BE148877		"gb:CM4-HT0244-111199-040-h12 HT0244 Hom	3.95
427318	AF186081	Hs.175783	zinc transporter	3.92
411762	AW860972		"gb:QV0-CT0387-180300-167-h07 CT0387 Hom	3.85
418668	AW407987	Hs.87150	Human clone A9A2BR11 (CAC)n/(GTG)n repea	3.75
458311	AF069478		"gb:AF069478 Homo sapiens astrocytoma li	3.61
403649				3.60
419682	H13139	Hs.92282	paired-like homeodomain transcription fa	3.58
412519	AA196241	Hs.73980	"troponin T1, skeletal, slow"	3.51
414206	AW276887	Hs.46609	ESTs	3.45
427419	NM_000200	Hs.177888	histatin 3	3.37
420777	AA280223	Hs.130865	ESTs	3.35
428134	AA421773	Hs.161008	ESTs	3.31
450218	R02018	Hs.168640	"Ank, mouse, homolog of"	3.30
433474	AI192195	Hs.147174	"EST, Highly similar to ubiquitin-protei	3.30
418833	AW974899	Hs.292776	ESTs	3.26
400440	X83957	Hs.83870	nebulin	3.16

	413778	AA090235	Hs.75535	"myosin, light polypeptide 2, regulatory	3.06
	423151	AW838068		"gb:QV3-LT0048-010300-109-f02 LT0048 Hom	3.05
	445060	AA830811	Hs.88808	ESTs	2.98
	457065	AI476318	Hs.192480	ESTs	2.95
5	432456	H00093		"gb:ph8f12u_19/1TV Outward Alu-primed hn	2.92
	405678				2.85
	406707	S73840	Hs.931	"myosin, heavy polypeptide 2, skeletal m	2.81
	444105	AW189097	Hs.166597	ESTs	2.78
	433968	AL157518	Hs.90421	PRO2463 protein	2.73
10	438522	AA809431	Hs.258886	ESTs	2.73
	436562	H71937	Hs.169756	"complement component 1, s subcomponent"	2.68
	412417	AA102268	Hs.42175	ESTs	2.67
	455590	BE072259		"gb:QV4-BT0536-271299-059-g04 BT0536 Hom	2.65
	415380	F07953	Hs.16085	*putative G-protein coupled receptor*	2.65
15	428729	AL162331	Hs.191436	hypothetical protein FLJ10619	2.64
	408537	AW207734		"gb:UI-H-BI2-age-h-01-0-UI.s1 NCI_CGAP_S	2.63
	424706	AA741336	Hs.152108	transcriptional unit N143	2.63
	413212	BE072092		"gb:PM4-BT0532-160200-003-b11 BT0532 Hom	2.63
	406704	M21665	Hs.929	"myosin, heavy polypeptide 7, cardiac mu	2.62
20	437507	AA758538	Hs.246882	ESTs	2.60
	410384	AI933794	Hs.42745	ESTs	2.58
	408074	R20723	Hs.124764	ESTs	2.58
	436653	AA829828	Hs.292402	ESTs	2.52
	458090	AI282149	Hs.56213	"ESTs, Highly similar to FXD3_HUMAN FORK	2.51
25	432003	AI689154	Hs.122972	ESTs	2.50
	436915	AA737400	Hs.142230	ESTs	2.50
	410028	AW576454	Hs.258553	ESTs	2.46
	448920	AW408009	Hs.22580	alkylglycerone phosphate synthase	2.45
	422046	AI638562		"gb:ts50a10.x1 NCI_CGAP_Ut1 Homo sapiens	2.44
30	451122	AA015767	Hs.193587	ESTs	2.40
	422646	H87863	Hs.151380	ESTs	2.36
	451237	AW600293		"gb:EST00049 pGEM-T library Homo sapiens	2.36
	400001			AFFX control: BioB-3	2.36
	415835	Z45365		"gb:HSC2NF061 normalized infant brain cD	2.36
35	439706	AW872527	Hs.59761	ESTs	2.36
	423341	AW242394	Hs.252495	ESTs	2.36
	436486	AA742221	Hs.120633	ESTs	2.35
	407449	AJ002784		gb:Homo sapiens mRNA; fetal brain cDNA 5	2.33
	430573	AA744550	Hs.136345	ESTs	2.32
40	401974				2.31
	443356	AL044498	Hs.133262	"ESTs, Weakly similar to PH0217 reverse	2.31
	430751	NM_012471	Hs.247868	transient receptor potential channel 5	2.25
	439128	AI949371	Hs.153089	ESTs	2.25
	448765	R15337	Hs.21958	"Homo sapiens cDNA FLJ10532 fis, clone N	2.25
45	451130	AI762250	Hs.211347	ESTs	2.24
	405420				2.23
	455029	AW851258		"gb:IL3-CT0220-160200-066-H06 CT0220 Hom	2.23
	438224	AA933999		"gb:on91f04.s1 Soares_NFL_T_GBC_S1 Homo	2.23
	407764	BE008347		"gb:CM0-BN0154-080400-325-h04 BN0154 Hom	2.23
50	413549	BE252470		"gb:601108292F1 NIH_MGC_16 Homo sapiens	2.23
	437010	AA741368	Hs.291434	ESTs	2.23
	435111	AI914279	Hs.213740	ESTs	2.22
	403375				2.21
	455060	AW853441		"gb:RC1-CT0252-030100-023-g09 CT0252 Hom	2.21
55	409792	AW854153		"gb:RC3-CT0254-060400-029-d03 CT0254 Hom	2.20
	421154	AA284333	Hs.267631	"Homo sapiens cDNA FLJ14269 fis, clone P	2.19
	401963				2.18
	435034	AF168711	Hs.159397	x 010 protein	2.18
	448996	AW998989	Hs.105749	*KIAA0553 protein*	2.18
60	436816	AW297599	Hs.255667	ESTs	2.17
	442252	AI733395	Hs.129124	ESTs	2.17
	419310	AA236233	Hs.188716	ESTs	2.16
	418579	H91800	Hs.124156	ESTs	2.16
	423315	R54109	Hs.26096	ESTs	2.16
65	432744	AA988835	Hs.38664	ESTs	2.15
	424492	AI133482	Hs.165210	ESTs	2.15
	424770	AA425562		"gb:zw46e05.r1 Soares_total_fetus_Nb2HF8	2.15
	437101	AA744518	Hs.120610	ESTs	2.15
	428793	AC004957	Hs.298975	"ESTs, Highly similar to collapsin-2-lik	2.15

	415708	H56475		"gb:yt87d11.r1 Soares_pineal_gland_N3HPG	2.13
	459619				2.12
	427506	AK000134	Hs.179100	hypothetical protein FLJ20127	2.12
	452508	AA804174	Hs.184354	ESTs	2.10
5	410881	AW809157		"gb:RC0-ST0118-041099-031-c07_1 ST0118 Homo sapiens cDNA, mRNA sequence"	2.10
	403087				2.10
	403869				2.10
	445028	D81194	Hs.282499	ESTs	2.10
	447884	H29505		"gb:ym60d10.r1 Soares infant brain 1NIB Homo sapiens cDNA clone 5', mRNA sequence"	2.10
10	414575	H11257	Hs.295233	ESTs	2.09
	420351	BE218221	Hs.190044	ESTs	2.08
	426998	BE274360		"gb:601121068F1 NIH_MGC_20 Homo sapiens cDNA clone 5', mRNA sequence"	2.08
	405455				2.08
	423843	AA332652		"gb:EST36627 Embryo, 8 week I Homo sapiens cDNA 5' end similar to similar to	
15				monoamine oxidase B, mRNA sequence"	2.08
	406135				2.07
	427046	BE246180	Hs.121385	ESTs	2.07
	403493				2.05
	444514	AI682905	Hs.270431	"ESTs, Weakly similar to ALU1_HUMAN ALU SUBFAMILY J SEQUENCE	
20				CONTAMINATION WARNING ENTRY [H.sapiens]"	2.05
	435884	AA701443	Hs.192868	ESTs	2.05
	419629	AB020695	Hs.91662	KIAA0888 protein	2.03
	405900				2.03
	457350	AW974438	Hs.194136	"ESTs, Moderately similar to AF091457 1 zinc finger protein RIN ZF [R.norvegicus]"	2.02
25	400007			AFFX control: BioDn-5	2.01
	406978	M64358		"gb:Human rhom-3 gene, exon."	2.00

TABLE 8A shows the accession numbers for those primekeys lacking a unigeneID in Table 8. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Pkey: Unique Eos probeset identifier number
CAT number: Gene cluster number
Accession: Genbank accession numbers

Pkey	CAT number	Accessions
407764	1014849_1	BE008347 BE008320 BE083307 BE083311 AW075968
408537	1064753_1	AW207734 D60164 D81150 D81078 D61356 AW996804
409792	1154677_1	AW854153 AW500210 BE145772 AW501310
410881	1225682_1	AW809157 AW812181 AW812175 AW812172 AW812161 AW812165
411762	1256906_1	AW860972 AW862598 AW862599 AW860988 AW860983 AW860898 AW860925 AW860922 AW860986 AW860984 AW860989
413212	1353792_1	BE072092 BE072106 BE072086 BE072098 BE072103
413549	1375933_2	BE252470 BE147573
415708	1548209_1	H56475 F29401 F34552
415835	1558511_1	Z45365 R25905 H05203 T77496
422046	210744_1	AI638562 T16929 H13401 F07773 R55836
423151	225415_1	AW838068 AW837986 AW838067 AA322487 AW837936
423843	232510_1	AA332652 AA331633 AW999369 AW902993 BE170475 AA378845 AW964175 AI475221
424770	243504_1	AA425562 AI880208 AA346646 N22655 AW811775 AW811786
426998	274259_-1	BE274360
432456	347718_2	H00093 H00079 H00070 H00054 H00049 H00063 AW905306 AW905241 AW905410 AW905307 AW905411 AW905240 AW905210 AW905352 AW905304 AW905239 AW905242 AW905243 H00087
438224	452656_1	AA933999 AA781181
447884	740749_1	H29505 R18575 Z43580 T48738 AI435454 BE004683
451237	863269_1	AW600293 AI767468
455029	1249374_1	AW851258 AW851435 AW851106 AW851421
455060	1251259_1	AW853441 BE145228 BE145218 BE145162 BE145283
455590	1335127_1	BE072259 BE072230 BE007911
458311	543550_1	AF069478 AF069479 AF069480

TABLE 8B shows the genomic positioning for those primekeys lacking unigene ID's and accession numbers in table 8. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

Pkey:	Unique number corresponding to an Eos probeset
Ref:	Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-495.
Strand:	Indicates DNA strand from which exons were predicted.
Nt_position:	Indicates nucleotide positions of predicted exons.

Pkey	Ref	Strand	Nt_position
401963	3126783	Plus	51382-51521
401974	3126777	Plus	85330-85683
403087	8954241	Plus	169511-169795
403375	9255944	Minus	92554-92795
403493	7341425	Plus	157568-159084
403612	8469060	Minus	94723-94859
403649	8705159	Minus	27141-27247
403869	7280046	Minus	34379-34583
404113	9588571	Minus	13446-13646
404567	7249169	Minus	101320-101501
405163	9966267	Minus	161171-161299
405227	6731245	Minus	22550-22802
405420	7211837	Minus	13428-13582
405455	7656675	Plus	134112-134671
405678	4079670	Plus	151821-152027
405900	6758795	Minus	71181-71535
406135	9164918	Minus	65489-65715

TABLE 9: 1001 GENES SIGNIFICANTLY UP-REGULATED IN NORMAL PROSTATE COMPATED TO PROSTATE CANCER

Table 9 shows 1001 genes significantly up-regulated in prostate cancer compared to normal prostate. These were selected from 59680 probesets on the Affymetrix/Eos Hu03 GeneChip array such that the ratio of "average" normal prostate to "average" prostate cancer tissues was greater than or equal to 8.14. The "average" normal prostate level was set to the mean amongst 4 normal prostate tissues. The "average" prostate cancer level was set to the 85th percentile amongst 73 tumor samples. In order to remove gene-specific background levels of non-specific hybridization, the 10th percentile value amongst all the tissues was subtracted from both the numerator and the denominator before the ratio was evaluated.

Pkey: Unique Eos probeset identifier number
ExAccn: Exemplar Accession number, Genbank accession number
UnigeneID: Unigene number
Unigene Title: Unigene gene title
R1: Ratio of prostate cancer to normal prostate

Pkey	ExAccn	UnigeneID	Unigene Title	R1
451002	AA013299	Hs.8018	ESTs, Weakly similar to ALU3_HUMAN ALU S	1684.00
435596	AA689465	Hs.188999	ESTs	738.00
443576	AI078027	Hs.169338	ESTs	246.86
434247	AA928116	Hs.272065	ESTs	245.20
400452	AK000185		gb:Homo sapiens cDNA FLJ20178 fis, clone	222.00
405932				221.33
427906	AA864330	Hs.166520	ESTs	212.00
443685	AI686550	Hs.174481	ESTs	163.20
451554	AI474866	Hs.193237	ESTs	149.45
418323	NM_002118	Hs.1162	major histocompatibility complex, class	126.11
429480	M36860	Hs.9295	elastin (supravalvular aortic stenosis,	123.27
426025	AW138330	Hs.233778	ESTs	120.00
418917	X02994	Hs.1217	adenosine deaminase	106.75
404407				105.71
442027	AI652926	Hs.128395	ESTs	100.53
433704	AA608684	Hs.121705	ESTs, Moderately similar to ALUC_HUMAN !	94.00
453758	U83527		gb:HSU83527 Human fetal brain (M.Lovett)	89.18
415354	F06495		gb:HSC1AB051 normalized infant brain cDN	87.73
424239	M67439	Hs.143526	dopamine receptor D5	86.82
444143	AW747996	Hs.160999	ESTs	86.43
401672				77.26
430590	AW383947	Hs.246381	CD68 antigen	68.47
411972	BE074959		gb:PM0-BT0582-310100-001-f08 BT0582 Homo	68.00
448992	AI766053	Hs.188346	ESTs	61.26
408828	BE540279		gb:601059857F1 NIH_MGC_10 Homo sapiens c	57.71
409653	AW451693	Hs.220826	ESTs	56.40
402964				54.67
422673	N59027		gb:yv59d11.r1 Soares fetal liver spleen	54.00
422568	AA372275	Hs.279800	Homo sapiens cDNA FLJ11383 fis, clone HE	54.00
438907	R32704	Hs.301298	ESTs	52.96
405172				52.96
444897	AW137088	Hs.144857	ESTs	52.32
458019	AW592931	Hs.256298	ESTs	51.63
405275	AB028989	Hs.88500	mitogen-activated protein kinase 8 inter	50.98
457815	AA703679	Hs.106999	ESTs, Weakly similar to SYT5_HUMAN SYNAP	49.60
424385	AA339666		gb:EST44776 Fetal brain I Homo sapiens c	48.90
407172	T54095		gb:ya92c05.s1 Stratagene placenta (93722	47.98
428202	AA424163	Hs.156895	ESTs	46.83
435672	AI700148	Hs.283626	ESTs	43.57
420283	AA485224	Hs.57734	G protein-coupled receptor kinase-intera	43.00
417016	AA837098	Hs.269933	ESTs	42.70
438854	AF074994	Hs.24240	ESTs	42.67

406134				42.43
457319	AA480895	Hs.201552	ESTs, Weakly similar to T17288 hypotheti	42.31
409314	AA070266		gb:zm69d04.r1 Stratagene neuroepithelium	42.25
401124				41.61
429316	AI371157	Hs.178538	ESTs	40.00
420317	AB006628	Hs.96485	KIAA0290 protein	39.64
457586	AW062439		gb:MR0-CT0060-120899-001-f08 CT0060 Homo	39.60
417407	AA923278	Hs.290905	ESTs, Weakly similar to protease [H.sapi	38.73
430269	BE221682	Hs.178364	ESTs	38.06
439602	W79114	Hs.58558	ESTs	36.69
433686	AA604799	Hs.136528	ESTs, Moderately similar to ALU1_HUMAN A	36.29
417993	AW963705	Hs.295806	ESTs, Weakly similar to ALU7_HUMAN ALU S	36.18
428214	AA936282	Hs.120397	ESTs	36.10
416908	AA333990	Hs.80424	coagulation factor XIII, A1 polypeptide	36.08
426264	BE314852	Hs.168694	hypothetical protein FLJ10257	36.00
415911	H08796	Hs.124952	ESTs	36.00
457502	AA076049	Hs.274415	Homo sapiens cDNA FLJ10229 fis, clone HE	35.23
421566	NM_000399	Hs.1395	early growth response 2 (Krox-20 (Drosop	35.20
401468				34.89
458561	AI220150	Hs.211195	ESTs	34.60
433601	BE350738	Hs.123993	ESTs, Weakly similar to T00366 hypotheti	33.24
454977	AW848032		gb:IL3-CT0214-231299-053-D11 CT0214 Homo	32.96
402828				32.93
414522	AW518944	Hs.76325	Homo sapiens cDNA: FLJ23125 fis, clone L	31.76
402842				31.68
421245	AA285363		gb:HTH280 HTCDL1 Homo sapiens cDNA 5'/3'	31.59
401631	F05183	Hs.1799	CD1D antigen, d polypeptide	31.26
408057	AW139565		gb:UI-H-BI1-aea-d-04-0-UI.s1 NCI_CGAP_Su	31.24
408069	H81795		gb:ys68a10.r1 Soares retina N2b4HR Homo	31.20
438694	T87479	Hs.291797	ESTs	31.09
449156	AF103907	Hs.171353	prostate cancer antigen 3	29.78
428796	AU076734	Hs.193665	solute carrier family 28 (sodium-coupled	29.76
452549	AI907039		gb:PM-BT134-020499-566 BT134 Homo sapien	29.59
410129	BE244074	Hs.285531	regulator of Fas-induced apoptosis	29.53
414464	AI870175	Hs.13957	ESTs	29.47
412326	R07566	Hs.73817	Small inducible cytokine A3 (homologous	29.22
459081	W07808		gb:zb03a12.r1 Soares_fetal_lung_NbHL19W	29.20
448702	AW102670	Hs.122464	ESTs	29.13
451939	U80456	Hs.27311	single-minded (Drosophila) homolog 2	28.74
443412	W84893	Hs.9305	angiotensin receptor-like 1	28.61
457324	AB028990	Hs.243901	KIAA1067 protein	28.24
424247	X14008	Hs.234734	lysozyme (renal amyloidosis)	28.18
457140	AI279960	Hs.178140	ESTs	28.12
444151	AW972917	Hs.128749	alpha-methylacyl-CoA racemase	28.06
457669	AW104257	Hs.123426	ESTs, Weakly similar to putative serine/	27.61
412429	AV650262	Hs.75765	GRO2 oncogene	27.36
405495				27.33
406516				27.25
407997	AW135429	Hs.243577	ESTs	26.96
442115	AW452332	Hs.257554	ESTs	26.36
409038	T97490	Hs.50002	small inducible cytokine subfamily A (Cy	26.34
402838				26.32
449846	AI979284	Hs.200552	ESTs	26.21
417153	X57010	Hs.81343	collagen, type II, alpha 1 (primary oste	26.20
439792	NM_014856	Hs.6684	KIAA0476 gene product	25.91
450096	AI682088	Hs.223368	ESTs	25.60
424196	AL133660	Hs.142926	Homo sapiens mRNA; cDNA DKFZp434M0927 (f	25.57
414246	BE391090	Hs.280278	EST	25.57
420848	NM_005188	Hs.99960	Cas-Br-M (murine) ecotropic retroviral t	25.48
424778	AA251048	Hs.153042	lymphocyte antigen 9	25.42
409126	AA063426		gb:zf70c08.s1 Soares_pineal_gland_N3HPG	25.25
443936	AW083491	Hs.31196	ESTs	25.22
419392	W28573		gb:51f10 Human retina cDNA randomly prim	25.01
411201	T74588	Hs.8509	ESTs, Weakly similar to CO3_HUMAN COMPLE	24.85
422940	BE077458		gb:RC1-BT0606-090500-015-b04 BT0606 Homo	24.76
437571	AA760894	Hs.153023	ESTs	24.74
433973	AI014723	Hs.131770	ESTs	24.57
422416	BE019557	Hs.11900	Human DNA sequence from clone RP4-583P15	24.53
421552	AF026692	Hs.105700	secreted frizzled-related protein 4	24.49

443668	U25758	Hs.134584	ESTs	24.49
424800	AL035588	Hs.153203	MyoD family inhibitor	24.10
453633	AA357001	Hs.34045	hypothetical protein FLJ20764	24.04
430565	AL122081	Hs.244343	cadherin related 23	24.00
433694	AI208611	Hs.12066	Homo sapiens cDNA FLJ11720 fis, clone HE	23.89
451045	AA215672		gb:zr96e09.s1 NCI_CGAP_GCB1 Homo sapiens	23.83
408583	AW449674	Hs.47359	ESTs	23.73
444040	AF204231	Hs.182982	golgin-67	23.62
414182	AA136301		gb:zk93g04.s1 Soares_pregnant_uterus_NbH	23.39
418678	NM_001327	Hs.167379	cancer/testis antigen	23.20
408380	AF123050	Hs.44532	diubiquitin	22.68
456076	BE243877	Hs.76941	ATPase, Na+/K+ transporting, beta 3 poly	22.65
418299	AA279530	Hs.83968	integrin, beta 2 (antigen CD18 (p95), ly	22.38
444917	R68651	Hs.144997	ESTs	22.26
444381	BE387335	Hs.283713	ESTs	22.08
415788	AW628686	Hs.78851	KIAA0217 protein	22.04
410896	AW809637		gb:MR4-ST0124-261099-015-b07 ST0124 Homo	22.00
412978	AI431708	Hs.820	homeo box C6	21.95
458418	AV653846	Hs.126261	Homo sapiens Chromosome 16 BAC clone CIT	21.94
454791	BE071874		gb:RC2-BT0522-120200-014-a06 BT0522 Homo	21.84
408748	J05500	Hs.47431	spectrin, beta, erythrocytic (includes s	21.26
416011	H14487		gb:ym18c10.r1 Soares infant brain 1NIB H	21.24
440474	AI207936	Hs.7195	gamma-aminobutyric acid (GABA) A recepto	21.14
447047	AI623698	Hs.246306	Homo sapiens cDNA: FLJ23529 fis, clone L	21.11
426793	X89887	Hs.172350	HIR (histone cell cycle regulation defec	21.10
409841	AW502139		gb:UI-HF-BR0p-ajr-e-05-0-UI.r1 NIH_MGC_5	21.07
405685				20.90
457359	AI983207	Hs.192481	ESTs, Weakly similar to SYPH_HUMAN SYNAP	20.84
423067	AA321355	Hs.285401	ESTs	20.74
422355	AW403724	Hs.140	immunoglobulin heavy constant gamma 3 (G	20.73
401201				20.73
458278	W28912	Hs.129019	ESTs	20.68
439097	H66948		gb:yr66d10.r1 Soares fetal liver spleen	20.67
414875	H42679	Hs.77522	major histocompatibility complex, class	20.66
400926				20.66
451355	NM_004197	Hs.444	serine/threonine kinase 19	20.64
446982	AW500221	Hs.43616	Homo sapiens mRNA for FLJ00029 protein,	20.61
417105	X60992	Hs.81226	CD6 antigen	20.61
405777				20.51
424123	AW966158	Hs.58582	Homo sapiens cDNA FLJ12702 fis, clone NT	20.20
425009	X58288	Hs.154151	protein tyrosine phosphatase, receptor t	20.10
443271	BE568568	Hs.195704	ESTs	19.98
421064	AI245432	Hs.101382	tumor necrosis factor, alpha-induced pro	19.98
418819	AA228776	Hs.191721	ESTs	19.94
457595	AA584854		gb:no09h11.s1 NCI_CGAP_Phe1 Homo sapiens	19.90
404426				19.84
412571	U43143	Hs.74049	fms-related tyrosine kinase 4	19.79
431457	NM_012211	Hs.256297	integrin, alpha 11	19.62
414002	NM_006732	Hs.75678	FBJ murine osteosarcoma viral oncogene h	19.57
418994	AA296520	Hs.89546	Selectin E (endothelial adhesion molecul	19.56
437158	AW090198	Hs.4779	KIAA1150 protein	19.52
437866	AA156781	Hs.83992	ESTs	19.44
417421	AL138201	Hs.82120	nuclear receptor subfamily 4, group A, m	19.34
433057	X15675	Hs.296832	Human pTR7 mRNA for repetitive sequence	19.22
421730	AW449808	Hs.164036	glucosamine (N-acetyl)-6-sulfatase (Sanf	19.21
456557	AA284477	Hs.96618	ESTs	18.77
440806	AI247422	Hs.129966	ESTs	18.76
439845	AL355743	Hs.56663	Homo sapiens EST from clone 41214, full	18.65
416155	AI807264	Hs.205442	ESTs, Weakly similar to AF117610 1 inner	18.64
437820	AA769062	Hs.16029	ESTs, Weakly similar to alternatively sp	18.62
450923	AW043951	Hs.38449	ESTs	18.59
418329	AW247430	Hs.84152	cystathionine-beta-synthase	18.58
424537	AI673027	Hs.143271	ESTs	18.55
447742	AF113925	Hs.19405	caspase recruitment domain 4	18.52
415251	R42863	Hs.7124	ESTs	18.47
440770	AA912815	Hs.222078	ESTs	18.40
407711	AI085846	Hs.25522	ESTs	18.32
427157	U51166	Hs.173824	thymine-DNA glycosylase	18.28
409847	AW501751	Hs.279733	ESTs	18.15

417240	N57568	Hs.176028	EST	18.13
435732	AF229178	Hs.123136	leucine rich repeat and death domain con	18.12
436896	AW977385	Hs.278615	ESTs	18.12
432485	N90866	Hs.276770	CDW52 antigen (CAMPATH-1 antigen)	17.90
429490	AI971131	Hs.293684	ESTs, Weakly similar to alternatively sp	17.82
429984	AL050102	Hs.227209	DKFZP586F1019 protein	17.82
449214	AI889114	Hs.195663	ESTs	17.75
433867	AK000596	Hs.3618	hippocalcin-like 1	17.72
431735	AW977724	Hs.75968	thymosin, beta 4, X chromosome	17.71
401515				17.67
444045	AI097439	Hs.135548	ESTs	17.58
442754	AL045825	Hs.210197	ESTs	17.55
426559	AB001914	Hs.170414	paired basic amino acid cleaving system	17.54
432415	T16971	Hs.289014	ESTs	17.50
427829	AI188225	Hs.127462	ESTs	17.50
432516	R08003	Hs.188013	ESTs	17.44
435259	AA152106	Hs.4859	cyclin L ania-6a	17.36
414989	T81668		gb:yd29c04.r1 Soares fetal liver spleen	17.31
444880	AW118683	Hs.154150	ESTs	17.30
417651	R06874	Hs.268628	ESTs	17.27
453457	AL037103	Hs.270599	ESTs, Weakly similar to unnamed protein	17.22
424246	AW452533	Hs.143604	Kaiso	17.22
419078	M93119	Hs.89584	insulinoma-associated 1	17.18
417696	BE241624	Hs.82401	CD69 antigen (p60, early T-cell activati	17.14
431117	AF003522	Hs.250500	delta (Drosophila)-like 1	17.14
455254	AW877015		gb:QV2-PT0010-250300-096-f12 PT0010 Homo	17.14
425782	U66468	Hs.159525	cell growth regulatory with EF-hand doma	17.12
426678	H08170	Hs.113755	ESTs	17.12
426403	NM_000361	Hs.2030	thrombomodulin	17.01
425905	AB032959	Hs.161700	KIAA1133 protein	17.00
438867	AW451157	Hs.181157	ESTs	16.98
420940	AA830664	Hs.143974	ESTs	16.94
459234	AI940425		gb:CM0-CT0052-150799-024-c04 CT0052 Homo	16.92
404756				16.91
422247	U18244	Hs.113602	solute carrier family 1 (high affinity a	16.90
420568	F09247	Hs.167399	protocadherin alpha 5	16.88
443559	AI076765	Hs.269899	ESTs	16.80
438703	AI803373	Hs.31599	ESTs	16.78
411424	AW845985		gb:RC2-CT0163-200999-002-H08 CT0163 Homo	16.70
402895				16.69
422538	NM_006441	Hs.118131	5,10-methenyltetrahydrofolate synthetase	16.68
447108	AW449602	Hs.217953	ESTs, Moderately similar to NK-TUMOR REC	16.65
448520	AB002367	Hs.21355	doublecortin and CaM kinase-like 1	16.54
438567	AW451955	Hs.153065	ESTs	16.52
407811	AW190902	Hs.40098	cysteine knot superfamily 1, BMP antagon	16.50
410721	R23534	Hs.2730	heterogeneous nuclear ribonucleoprotein	16.50
437133	AB018319	Hs.5460	KIAA0776 protein	16.40
408182	AA047854		gb:zf49g04.r1 Soares retina N2b4HR Homo	16.32
417315	AI080042	Hs.180450	ribosomal protein S24	16.30
431840	AA534908	Hs.2860	POU domain, class 5, transcription facto	16.28
439882	AA847856	Hs.124565	ESTs	16.20
418277	AW135221	Hs.130812	ESTs	16.09
410688	AW796342		gb:PM2-UM0027-230200-002-h02 UM0027 Homo	16.04
420120	AL049610	Hs.95243	transcription elongation factor A (SII)-	16.04
429597	NM_003816	Hs.2442	a disintegrin and metalloproteinase doma	16.02
447033	AI357412	Hs.157601	EST - not in UniGene	16.02
421684	BE281591	Hs.106768	hypothetical protein FLJ10511	15.94
408599	AA055800	Hs.222933	ESTs	15.93
446012	AV656098	Hs.172382	hypothetical protein FLJ20001	15.86
409671	AA076769		gb:7B02B10 Chromosome 7 Fetal Brain cDNA	15.85
405934				15.84
426108	AA622037	Hs.166468	programmed cell death 5	15.84
416208	AW291168	Hs.41295	ESTs	15.48
410708	AA534370	Hs.154088	Homo sapiens cDNA: FLJ22756 fis, clone K	15.42
447342	AI199268	Hs.19322	ESTs; Weakly similar to !!!! ALU SUBFAMI	15.38
454563	AW807530		gb:CM0-ST0081-130999-054-d02 ST0081 Homo	15.37
411507	AW850140		gb:IL3-CT0219-261099-023-D11 CT0219 Homo	15.36
438170	AI916685	Hs.194601	ESTs	15.29
416292	AA179233	Hs.42390	nasopharyngeal carcinoma susceptibility	15.26

406638	M13861		gb:Human T-cell receptor active beta-cha	15.26
446686	AW138043	Hs.156307	ESTs	15.25
434485	AI623511	Hs.118567	ESTs	15.24
441188	AW292830	Hs.255609	ESTs	15.22
444172	BE147740	Hs.104558	ESTs	15.22
409521	BE244854	Hs.159578	Homo sapiens mRNA for FLJ00020 protein,	15.16
420748	AA279956	Hs.88672	ESTs	15.14
422583	AA410506	Hs.118578	H.sapiens mRNA for ribosomal protein L18	15.14
424240	AB023185	Hs.143535	calcium/calmodulin-dependent protein kin	15.12
451118	AI862096	Hs.60640	ESTs	15.12
437495	BE177778		gb:RC1-HT0598-310300-012-f07 HT0598 Homo	15.12
445467	AI239832	Hs.15617	ESTs, Weakly similar to ALU4_HUMAN ALU S	15.06
418305	AW006783	Hs.6686	ESTs	15.03
402812				15.02
436851	AA732480	Hs.293581	ESTs	15.00
400991				15.00
415752	BE314524	Hs.78776	Human putative transmembrane protein (nm	14.96
429900	AA460421	Hs.30875	ESTs	14.90
403683				14.84
430315	NM_004293	Hs.239147	guanine deaminase	14.80
451952	AL120173	Hs.301663	ESTs	14.72
424687	J05070	Hs.151738	matrix metalloproteinase 9 (gelatinase B	14.69
447229	BE617135		gb:601441677F1 NIH_MGC_65 Homo sapiens c	14.67
425818	AB021225	Hs.159581	matrix metalloproteinase 17 (membrane-in	14.65
448553	AI638449	Hs.173031	ESTs	14.63
431089	BE041395	Hs.283676	ESTs, Weakly similar to unknown protein	14.60
459145	AI903354		gb:RC-BT029-100199-117 BT029 Homo sapien	14.55
449650	AF055575	Hs.297647	ESTs, Moderately similar to calcium chan	14.54
400952				14.46
445885	AI734009	Hs.127699	EST cluster (not in UniGene)	14.44
407938	AA905097	Hs.85050	phospholamban	14.42
431676	AI685464	Hs.292638	ESTs	14.40
437210	AA311443	Hs.293563	Homo sapiens mRNA; cDNA DKFZp586E2317 (f	14.36
451900	AB023199	Hs.27207	KIAA0982 protein	14.36
445800	AA126419	Hs.301632	ESTs	14.32
412368	AW945992	Hs.181125	immunoglobulin lambda locus	14.31
409055	AW304028	Hs.300578	ESTs	14.23
408763	W57550	Hs.301526	Homo sapiens cDNA FLJ13181 fis, clone NT	14.22
446734	AL049278	Hs.16074	Homo sapiens mRNA; cDNA DKFZp564I153 (fr	14.22
413551	BE242639	Hs.75425	ubiquitin associated protein	14.22
421913	AI934365	Hs.109439	osteoglycin (osteoinductive factor, mime	14.22
452712	AW838616		gb:RC5-LT0054-140200-013-D01 LT0054 Homo	14.22
451468	AW503398	Hs.210047	ESTs	14.16
406038	Y14443	Hs.88219	zinc finger protein 200	14.14
424909	S78187	Hs.153752	cell division cycle 25B	14.07
434078	AW880709	Hs.283683	EST	14.07
415254	AI815831	Hs.184378	ESTs	14.05
418196	AI745649	Hs.26549	ESTs, Weakly similar to T00066 hypotheti	14.02
410020	T86315	Hs.728	ribonuclease, RNase A family, 2 (liver,	13.98
411352	NM_002890	Hs.758	RAS p21 protein activator (GTPase activa	13.98
429848	AF145439	Hs.225946	chemokine (C-C motif) receptor 9	13.95
413729	BE159999		gb:QV1-HT0412-270300-123-d10 HT0412 Homo	13.90
400125				13.88
420319	AW406289	Hs.96593	hypothetical protein	13.85
448272	AI479094	Hs.170786	ESTs	13.80
422695	AA315158		gb:EST186956 HCC cell line (matastasis t	13.80
424565	AW102723	Hs.75295	guanylate cyclase 1, soluble, alpha 3	13.78
458048	H30340	Hs.173705	Homo sapiens cDNA: FLJ22050 fis, clone H	13.78
408894	AI935400	Hs.217286	ESTs	13.76
454093	AW860158		gb:RC0-CT0379-290100-032-b04 CT0379 Homo	13.75
410889	X91662	Hs.66744	twist (Drosophila) homolog (acrocephalos	13.74
457751	AI908236		gb:IL-BT166-180399-010 BT166 Homo sapien	13.72
455131	AW857913		gb:RC0-CT0323-231199-031-b05 CT0323 Homo	13.69
408364	AW015238	Hs.128453	ESTs	13.67
425907	AA365752	Hs.155965	ESTs	13.62
402359				13.60
401044				13.53
409877	AW502498	Hs.157150	ESTs, Weakly similar to zinc finger prot	13.53
423690	AA329648	Hs.23804	ESTs	13.49

	430685	AI690234	Hs.191666	ESTs, Weakly similar to reverse transcri	13.47
	414052	AW578849	Hs.283552	ESTs, Weakly similar to unnamed protein	13.46
	447858	AW080339	Hs.211911	ESTs	13.44
	435716	AI573283	Hs.38458	ESTs	13.44
5	439120	H56389		gb:yt87c03.r1 Soares_pineal_gland_N3HPG	13.43
	402788				13.40
	451591	AA886446	Hs.146278	ESTs	13.40
	405411				13.38
	426558	AW188574	Hs.24218	ESTs	13.34
10	453506	AA132818	Hs.110407	ESTs, Weakly similar to coded for by C.	13.33
	416445	AL043004	Hs.300678	Human serine/threonine kinase mRNA, part	13.32
	457084	AI074149	Hs.150905	ESTs, Weakly similar to chondroitin 4-su	13.32
	403838				13.32
	427337	Z46223	Hs.176663	Fc fragment of IgG, low affinity IIIb, r	13.30
15	434318	AW207552	Hs.116328	ESTs, Weakly similar to dJ134E15.1 [H.sa	13.28
	435193	N41359	Hs.218107	ESTs	13.28
	414756	AW451101	Hs.159489	ESTs, Moderately similar to hexokinase I	13.27
	420626	AF043722	Hs.99491	RAS guanyl releasing protein 2 (calcium	13.26
	420052	AA418850	Hs.44410	ESTs	13.25
20	414020	NM_002984	Hs.75703	small inducible cytokine A4 (homologous	13.25
	403851				13.24
	422647	W07492	Hs.157101	ESTs	13.21
	433598	AI762836	Hs.271433	ESTs, Moderately similar to ALU2_HUMAN A	13.21
	409065	AB033113	Hs.50187	KIAA1287 protein	13.20
25	435063	R21966	Hs.57734	G protein-coupled receptor kinase-intera	13.19
	439367	BE386844	Hs.248746	ESTs	13.17
	451957	AI796320	Hs.10299	Homo sapiens cDNA FLJ13545 fis, clone PL	13.16
	420569	AA278362	Hs.289062	Homo sapiens cDNA FLJ12334 fis, clone MA	13.14
	447883	BE262802	Hs.4909	dickkopf (Xenopus laevis) homolog 3	13.07
30	426490	NM_001621	Hs.170087	aryl hydrocarbon receptor	13.06
	414789	AA155859	Hs.79708	ESTs	13.05
	451418	BE387790	Hs.26369	ESTs	13.04
	443494	T99719	Hs.270404	Homo sapiens cDNA: FLJ22389 fis, clone H	13.03
	425878	AW964806	Hs.38085	ESTs, Weakly similar to putative glycine	13.02
35	431912	AI660552	Hs.154903	ESTs, Weakly similar to A56154 Abl subst	13.00
	407122	H20276	Hs.31742	ESTs	13.00
	456491	AL137466	Hs.97277	Homo sapiens mRNA; cDNA DKFZp434H1322 (f	12.99
	448172	N75276	Hs.135904	ESTs	12.98
	452144	AA032197	Hs.102558	ESTs	12.96
40	419953	BE267154	Hs.125752	ESTs	12.96
	416182	NM_004354	Hs.79069	cyclin G2	12.94
	451154	AA015879	Hs.33536	ESTs	12.93
	412257	AW903830		gb:CM4-NN1037-250400-155-h04 NN1037 Homo	12.93
	449784	AW161319	Hs.12915	ESTs	12.92
45	432695	D63480	Hs.278634	KIAA0146 protein	12.92
	454105	NM_001259	Hs.38481	cyclin-dependent kinase 6	12.92
	439093	AA534163	Hs.5476	serine protease inhibitor, Kazal type, 5	12.90
	416098	H41324	Hs.31581	ESTs, Moderately similar to ST1B_HUMAN S	12.88
	424897	D63216	Hs.153684	frizzled-related protein	12.88
50	414604	AU076649	Hs.76556	growth arrest and DNA-damage-inducible 3	12.88
	414664	AA587775	Hs.66295	Homo sapiens HSPC311 mRNA, partial cds	12.84
	452560	BE077084		gb:RC5-BT0603-220200-013-C07 BT0603 Homo	12.84
	413869	NM_000878	Hs.75596	interleukin 2 receptor, beta	12.80
	452359	BE167229	Hs.29206	Homo sapiens clone 24659 mRNA sequence	12.80
55	435886	BE265839	Hs.12126	hepatocellular carcinoma-associated anti	12.78
	445230	U97018	Hs.12451	echinoderm microtubule-associated protei	12.78
	412226	W26786		gb:15d7 Human retina cDNA randomly prime	12.77
	446619	AU076643	Hs.313	secreted phosphoprotein 1 (osteopontin,	12.76
	447769	AW873704	Hs.48764	ESTs	12.76
60	414478	AI306389	Hs.76240	adenylate kinase 1	12.76
	425383	D83407	Hs.156007	Down syndrome critical region gene 1-lik	12.68
	450704	H85157	Hs.40696	ESTs	12.66
	405856				12.66
	412935	BE267045	Hs.75064	tubulin-specific chaperone c	12.65
65	402802				12.62
	452588	AA889120	Hs.110637	Homeo box A10	12.62
	419978	NM_001454	Hs.93974	forkhead box J1	12.62
	403137				12.60
	430226	BE245562	Hs.2551	adrenergic, beta-2-, receptor, surface	12.57

448076	AJ133123	Hs.20196	adenylate cyclase 9	12.56
450462	F07097	Hs.300828	Homo sapiens mRNA full length insert cDN	12.54
405236				12.52
409292	AA071051		gb:zm58e05.s1 Stratagene fibroblast (937	12.47
421540	AA767669	Hs.10242	ESTs	12.47
425840	AW978731	Hs.301824	ESTs	12.44
443181	AI039201	Hs.54548	ESTs	12.42
452436	BE077546	Hs.31447	ESTs	12.42
455183	AW984111		gb:RC0-HN0007-160300-011-f09 HN0007 Homo	12.40
432887	AI926047	Hs.162859	ESTs	12.37
410494	M36564	Hs.64016	protein S (alpha)	12.36
439024	R96696	Hs.35598	ESTs	12.36
451246	AW189232	Hs.39140	cutaneous T-cell lymphoma tumor antigen	12.36
432892	AL042615	Hs.15995	ESTs	12.35
418982	AI348838	Hs.13073	ESTs	12.35
414516	AI307802	Hs.279551	ESTs	12.34
440134	BE410734		gb:601301619F1 NIH_MGC_21 Homo sapiens c	12.29
443873	AL048542	Hs.16291	ESTs	12.28
401286				12.26
454020	AW962845	Hs.256527	ESTs	12.24
420077	AW512260	Hs.87767	ESTs	12.24
443837	AI984625	Hs.9884	spindle pole body protein	12.24
407519	X64979		gb:H.sapiens mRNA HTPCRX01 for olfactory	12.23
435839	AF249744	Hs.25951	Rho guanine nucleotide exchange factor (	12.22
448552	AW973653	Hs.20104	hypothetical protein FLJ00052	12.20
405325				12.20
451009	AA013140	Hs.115707	ESTs	12.18
423066	Y18264	Hs.120171	ESTs	12.17
439556	AI623752	Hs.163603	ESTs	12.16
443062	N77999	Hs.8963	Homo sapiens mRNA full length insert cDN	12.15
445873	AA250970	Hs.251946	Homo sapiens cDNA: FLJ23107 fis, clone L	12.14
453542	AW836724	Hs.33190	Homo sapiens mRNA expressed only in plac	12.11
440106	AA864968	Hs.127699	ESTs	12.10
417605	AF006609	Hs.82294	regulator of G-protein signalling 3	12.10
440286	U29589	Hs.7138	cholinergic receptor, muscarinic 3	12.04
420061	AW024937	Hs.29410	ESTs	12.02
458727	AI022813	Hs.92679	Homo sapiens clone CDABP0014 mRNA sequen	11.96
445407	AI222658	Hs.221889	ESTs, Weakly similar to la costa [D.mela	11.95
418250	U29926	Hs.83918	adenosine monophosphate deaminase (isofo	11.94
414129	AI990287	Hs.270798	ESTs	11.93
409799	D11928	Hs.76845	phosphoserine phosphatase-like	11.92
438461	AW075485	Hs.286049	phosphoserine aminotransferase	11.92
443912	R37257	Hs.184780	ESTs	11.92
424606	AA343936		gb:EST49786 Gall bladder I Homo sapiens	11.90
434217	AW014795	Hs.23349	ESTs	11.90
451533	NM_004657	Hs.26530	serum deprivation response (phosphatidyl	11.90
422423	AF283777	Hs.116481	CD72 antigen	11.89
409398	AW386461		gb:PM4-PT0019-121299-004-F02 PT0019 Homo	11.89
423853	AB011537	Hs.133466	slit (Drosophila) homolog 1	11.82
446180	AI074413	Hs.14220	hypothetical protein FLJ20450	11.80
414341	D80004	Hs.75909	KIAA0182 protein	11.80
406538				11.79
433253	AW450502	Hs.24218	ESTs	11.79
447397	BE247676	Hs.18442	E-1 enzyme	11.78
451684	AF216751	Hs.26813	CDA14	11.76
416862	R23765	Hs.23575	ESTs	11.74
425770	NM_014363	Hs.159492	spastic ataxia of Charlevoix-Saguenay (s	11.72
428826	AL048842	Hs.194019	attractin	11.72
433037	NM_014158	Hs.279938	HSPC067 protein	11.72
447476	BE293466	Hs.20880	ESTs	11.72
452092	BE245374	Hs.27842	hypothetical protein FLJ11210	11.72
412922	M60721	Hs.74870	H2.0 (Drosophila)-like homeo box 1	11.72
401680	NM_005578	Hs.180398	LIM domain-containing preferred transloc	11.69
422576	BE548555	Hs.118554	CGI-83 protein	11.68
450203	AF097994	Hs.301528	L-kynurenine/alpha-aminoadipate aminotra	11.68
410531	AW752953		gb:QV0-CT0224-261099-035-g02 CT0224 Homo	11.67
425917	W28517	Hs.117167	Homo sapiens cDNA: FLJ23067 fis, clone L	11.66
418693	AI750878	Hs.87409	thrombospondin 1	11.64
400557				11.62

416188	BE157260	Hs.79070	v-myc avian myelocytomatosis viral oncog	11.60
419047	AW952771	Hs.90043	ESTs	11.59
420441	AI986160	Hs.88446	ESTs	11.59
400885				11.57
409853	AW502327		gb:UI-HF-BR0p-aka-a-07-0-UI.r1 NIH_MGC_5	11.56
400802				11.56
434540	NM_016045	Hs.5184	TH1 drosophila homolog	11.55
431449	M55994	Hs.256278	tumor necrosis factor receptor superfami	11.55
425928	S55736	Hs.238852	ESTs, Weakly similar to hypothetical pro	11.54
434701	AA460479	Hs.4096	KIAA0742 protein	11.53
434228	Z42047	Hs.283978	ESTs; KIAA0738 gene product	11.52
420729	AW964897	Hs.290825	ESTs	11.52
428328	AA426080	Hs.98489	ESTs	11.50
433887	AW204232	Hs.279522	ESTs	11.50
414812	X72755	Hs.77367	monokine induced by gamma interferon	11.46
457718	F18572	Hs.22978	ESTs	11.44
452260	AA453208	Hs.28726	RAB9, member RAS oncogene family	11.42
459029	AA131376	Hs.285203	fibroblast growth factor 12	11.42
456267	AI127958	Hs.83393	cystatin E/M	11.39
433285	AW975944	Hs.237396	ESTs	11.38
449186	AW291876	Hs.196986	ESTs	11.37
447861	AI434593	Hs.164294	ESTs	11.37
456023	R00028		gb:ye70a06.s1 Soares fetal liver spleen	11.36
439444	AI277652	Hs.54578	ESTs	11.31
401163				11.31
430886	L36149	Hs.248116	chemokine (C motif) XC receptor 1	11.28
450784	AW246803	Hs.47289	ESTs	11.28
452391	AL044829	Hs.29331	carnitine palmitoyltransferase I, muscle	11.27
449625	NM_014253	Hs.23796	odz (odd Oz/ten-m, Drosophila) homolog 1	11.26
456827	AA075687	Hs.147176	epidermal growth factor receptor substra	11.24
439328	W07411	Hs.118212	ESTs, Moderately similar to ALU3_HUMAN A	11.24
432093	H28383		gb:yl52c03.r1 Soares breast 3NbHBst Homo	11.24
407335	AA631047	Hs.158761	Homo sapiens cDNA FLJ13054 fis, clone NT	11.23
442501	AA315267	Hs.23128	ESTs	11.22
429746	AJ237672	Hs.214142	5,10-methylenetetrahydrofolate reductase	11.21
422858	R35398		gb:yg64g10.r1 Soares infant brain 1NIB H	11.20
415156	X84908	Hs.78060	phosphorylase kinase, beta	11.20
446713	AV660122	Hs.282675	ESTs	11.20
452221	C21322	Hs.11577	ESTs	11.20
418261	W78902	Hs.293297	ESTs	11.17
433332	AI367347	Hs.127809	ESTs	11.16
434539	AW748078	Hs.214410	ESTs	11.16
413471	BE142098		gb:CM4-HT0137-220999-017-d11 HT0137 Homo	11.14
410037	AB020725	Hs.58009	KIAA0918 protein	11.14
405601				11.13
458332	AI000341	Hs.220491	ESTs	11.12
427654	AA410183	Hs.137475	ESTs	11.12
427138	N77624	Hs.173717	phosphatidic acid phosphatase type 2B	11.10
431475	AI567669	Hs.287316	ESTs	11.10
425710	AF030880	Hs.159275	solute carrier family, member 4	11.08
413748	AW104057	Hs.19193	ESTs	11.07
409208	Y00093	Hs.51077	integrin, alpha X (antigen CD11C (p150),	11.07
457278	W92745	Hs.193324	ESTs	11.03
407021	U52077		gb:Human mariner1 transposase gene, comp	11.02
445701	AF055581	Hs.13131	lymphocyte adaptor protein	11.02
408338	AW867079		gb:MR1-SN0033-120400-002-c10 SN0033 Homo	10.95
401030	BE382701	Hs.25960	v-myc avian myelocytomatosis viral relat	10.95
437891	AW006969	Hs.6311	hypothetical protein FLJ20859	10.94
453874	AW591783	Hs.36131	collagen, type XIV, alpha 1 (undulin)	10.94
421562	AA530994	Hs.105803	ghrelin precursor	10.92
413431	AW246428	Hs.75355	ubiquitin-conjugating enzyme E2N (homolo	10.92
400132				10.92
436420	AA443966	Hs.31595	ESTs	10.90
424880	NM_000328	Hs.153614	retinitis pigmentosa GTPase regulator	10.88
433264	D85782	Hs.3229	cysteine dioxygenase, type I	10.88
429842	AI366213	Hs.173422	KIAA1605 protein	10.87
412405	AW948126		gb:RC0-MT0013-280300-031-a12 MT0013 Homo	10.85
400615				10.80
425018	BE245277	Hs.154196	E4F transcription factor 1	10.80

456011	BE243628		gb:TCBAP1D1053 Pediatric pre-B cell acut	10.79
455982	BE176862		gb:RC4-HT0587-170300-012-a04 HT0587 Homo	10.74
450418	BE218418	Hs.201802	ESTs	10.73
412490	AW803564	Hs.288850	ESTs	10.72
436962	AW377314	Hs.5364	DKFZP564I052 protein	10.70
437743	AI383497	Hs.131811	ESTs, Weakly similar to ALU1_HUMAN ALU S	10.70
449967	R40978	Hs.271498	ESTs, Moderately similar to ALU1_HUMAN A	10.70
449590	AA694070	Hs.268835	ESTs	10.68
446035	NM_006558	Hs.13565	Sam68-like phosphotyrosine protein, T-ST	10.68
426530	U24578	Hs.170250	complement component 4A	10.66
428600	AW863261	Hs.15036	ESTs, Highly similar to AF161358 1 HSPC0	10.64
420090	AA220238	Hs.94986	ribonuclease P (38kD)	10.64
451593	AF151879	Hs.26706	CGI-121 protein	10.62
438893	AF075031	Hs.29327	ESTs	10.62
459324	AW080953		gb:xc28c12.x1 NCI_CGAP_Co18 Homo sapiens	10.61
439883	AL359652	Hs.171096	Homo sapiens EST from clone DKFZp434A041	10.58
406513	AA715328	Hs.291205	ESTs	10.57
407826	AA128423	Hs.40300	calpain 3, (p94)	10.57
419550	D50918	Hs.90998	KIAA0128 protein; septin 2	10.56
428522	R10184	Hs.191987	ESTs, Weakly similar to ALU1_HUMAN ALU S	10.56
459526	AI142350	Hs.146735	EST	10.55
411448	AA178955	Hs.271439	ESTs	10.54
410102	AW248508	Hs.279727	ESTs;	10.52
406577				10.52
408405	AK001332	Hs.44672	hypothetical protein FLJ10470	10.51
428966	AF059214	Hs.194687	cholesterol 25-hydroxylase	10.50
400880				10.48
415875	AA894876	Hs.5687	protein phosphatase 1B (formerly 2C), ma	10.48
434715	BE005346	Hs.116410	ESTs	10.46
406851	AA609784	Hs.180255	major histocompatibility complex, class	10.44
413409	AI638418	Hs.21745	ESTs	10.44
418469	U76421	Hs.85302	adenosine deaminase, RNA-specific, B1 (h	10.44
419465	AW500239	Hs.21187	Homo sapiens cDNA: FLJ23068 fis, clone L	10.44
419544	AI909154		gb:QV-BT200-010499-007 BT200 Homo sapien	10.44
432180	Y18418	Hs.272822	RuvB (E coli homolog)-like 1	10.44
413822	R08950	Hs.272044	ESTs, Weakly similar to ALU1_HUMAN ALU S	10.42
437446	AA788946	Hs.16869	ESTs, Moderately similar to CA1C RAT COL	10.41
415701	NM_003878	Hs.78619	gamma-glutamyl hydrolase (conjugase, fol	10.41
443790	NM_003500	Hs.9795	acyl-Coenzyme A oxidase 2, branched chai	10.40
458873	AW150717	Hs.296176	STAT induced STAT inhibitor 3	10.38
415082	AA160000	Hs.137396	ESTs	10.37
429124	AW505086	Hs.196914	minor histocompatibility antigen HA-1	10.36
417187	AB011151	Hs.81505	KIAA0579 protein	10.34
426827	AW067605	Hs.172665	methylenetetrahydrofolate dehydrogenase	10.34
424280	NM_000030	Hs.271366	alanine-glyoxylate aminotransferase homo	10.33
446099	T93096	Hs.17126	ESTs	10.32
423445	NM_014324	Hs.128749	alpha-methylacyl-CoA racemase	10.31
409995	AW960597	Hs.30164	ESTs	10.30
432242	AW022715	Hs.162160	ESTs, Weakly similar to ALU4_HUMAN ALU S	10.30
406394	AA172106	Hs.110950	Rag C protein	10.30
406189				10.29
422283	AW411307	Hs.114311	CDC45 (cell division cycle 45, S.cerevis	10.26
401598	AA172106	Hs.110950	Rag C protein	10.26
456995	T89832	Hs.170278	ESTs	10.26
416511	NM_006762	Hs.79356	Lysosomal-associated multispanning membr	10.24
427274	NM_005211	Hs.174142	colony stimulating factor 1 receptor, fo	10.24
401384				10.23
456226	D13168	Hs.82002	endothelin receptor type B	10.22
426928	AF037062	Hs.172914	retinol dehydrogenase 5 (11-cisand 9-cis	10.21
423032	AI684746	Hs.119274	ESTs	10.20
436556	AI364997	Hs.7572	ESTs	10.20
418400	BE243026	Hs.301989	KIAA0246 protein	10.19
437401	AA757196	Hs.121190	ESTs	10.19
403690				10.17
423790	BE152393		gb:CM2-HT0323-171199-033-a08 HT0323 Homo	10.16
434094	AA305599	Hs.238205	hypothetical protein PRO2013	10.16
434967	AW975009	Hs.292274	ESTs	10.16
432827	Z68128	Hs.3109	Rho GTPase activating protein 4	10.16
432660	AI288430	Hs.64004	ESTs	10.14

452234	AW084176	Hs.223296	ESTs	10.14
445629	AI245701		gb:qk31f05.x1 NCI_CGAP_Kid3 Homo sapiens	10.13
457236	AA626142	Hs.179991	ESTs, Weakly similar to KPCE_HUMAN PROTE	10.13
444605	AI174603	Hs.254105	enolase 1, (alpha)	10.12
450313	AI038989	Hs.24809	hypothetical protein FLJ10826	10.12
407482	NM_006056			10.12
449971	AA807346	Hs.288581	Homo sapiens cDNA FLJ14296 fis, clone PL	10.11
441201	AW118822	Hs.128757	ESTs	10.10
435157	AW014605	Hs.179872	ESTs	10.10
417308	H60720	Hs.81892	KIAA0101 gene product	10.09
442582	AI204266	Hs.179303	ESTs	10.05
437252	AI433833	Hs.164159	ESTs, Weakly similar to ALU1_HUMAN ALU S	10.04
448663	BE614599	Hs.106823	H.sapiens gene from PAC 42616, similar t	10.04
434467	BE552368	Hs.231853	Homo sapiens cDNA FLJ13445 fis, clone PL	10.04
423698	AA329796	Hs.1098	DKFZp434J1813 protein	10.02
412707	AW206373	Hs.16443	Homo sapiens cDNA: FLJ21721 fis, clone C	10.00
414658	X58528	Hs.76781	ATP-binding cassette, sub-family D (ALD)	10.00
421832	NM_016098	Hs.108725	HSPC040 protein	10.00
423554	M90516	Hs.1674	glutamine-fructose-6-phosphate transamin	10.00
452039	AI922988	Hs.172510	ESTs	10.00
434673	AW137442	Hs.136965	ESTs	10.00
427978	AA418280	Hs.180040	Homo sapiens cDNA: FLJ22439 fis, clone H	10.00
457803	BE501815	Hs.198011	ESTs	9.99
428279	AA425310	Hs.155766	ESTs	9.98
444412	AI147652	Hs.216381	Homo sapiens clone HH409 unknown mRNA	9.98
417049	N72394	Hs.44862	ESTs	9.96
427509	M62505	Hs.2161	complement component 5 receptor 1 (C5a l	9.96
445424	AB028945	Hs.12696	cortactin SH3 domain-binding protein	9.96
443678	AW009605	Hs.231923	ESTs	9.96
447567	AW474513	Hs.224397	ESTs, Weakly similar to B48013 proline-r	9.94
414709	AA704703	Hs.77031	Sp2 transcription factor	9.94
434596	T59538		gb:yb65g12.s1 Stratagene ovary (937217)	9.94
427630	BE276115	Hs.144980	ESTs, Weakly similar to CA13_HUMAN COLLA	9.93
416111	AA033813	Hs.79018	chromatin assembly factor 1, subunit A (	9.92
423349	AF010258	Hs.127428	homeo box A9	9.92
424308	AW975531	Hs.154443	minichromosome maintenance deficient (S.	9.92
416814	AW192307	Hs.80042	dolichyl-P-Glc:Man9GlcNAc2-PP-dolichylgl	9.90
417986	AA481003	Hs.97128	ESTs	9.90
425174	D87450	Hs.154978	KIAA0261 protein	9.90
438171	AW976507	Hs.293515	ESTs	9.90
421984	AW972187	Hs.110443	hypothetical protein FLJ22215	9.89
408597	NM_005291	Hs.46453	G protein-coupled receptor 17	9.88
413907	AI097570	Hs.71222	ESTs	9.87
451296	AW801383	Hs.118578	H.sapiens mRNA for ribosomal protein L18	9.86
433409	AI278802	Hs.25661	ESTs	9.85
450360	AW117416	Hs.245484	ESTs	9.85
433104	AL043002	Hs.128246	ESTs, Moderately similar to unnamed prot	9.84
449824	AI962552	Hs.226765	ESTs	9.84
452744	AI267652	Hs.30504	Homo sapiens mRNA; cDNA DKFZp434E082 (fr	9.82
431066	AF026273	Hs.249175	interleukin-1 receptor-associated kinase	9.82
426457	AW894667	Hs.169965	chimerin (chimaerin) 1	9.80
443371	AI792868	Hs.145489	ESTs	9.80
437159	AL050072		gb:Homo sapiens mRNA; cDNA DKFZp566E1346	9.75
425242	D13635	Hs.155287	KIAA0010 gene product	9.74
447498	N67619	Hs.43687	ESTs	9.74
426759	AI590401	Hs.21213	ESTs	9.73
435129	AI381659	Hs.267086	ESTs	9.72
437672	AW748265	Hs.5741	flavohemoprotein b5+b5R	9.72
438209	AL120659	Hs.6111	KIAA0307 gene product	9.72
438440	AA807228	Hs.225161	ESTs	9.72
449720	AA311152	Hs.288708	ESTs; Weakly similar to KIAA0226 [H.sapi	9.72
414291	AI289619	Hs.13040	ESTs	9.72
436206	AK001451	Hs.265561	CD2-associated protein	9.70
446896	T15767	Hs.22452	Homo sapiens cDNA: FLJ21084 fis, clone C	9.70
412667	AW977540	Hs.269254	ESTs	9.70
423301	S67580	Hs.1645	cytochrome P450, subfamily IVA, polypept	9.67
440757	AW118645	Hs.160004	ESTs	9.67
441412	AI393657	Hs.159750	ESTs	9.66
421044	AF061871	Hs.101302	collagen, type XII, alpha 1	9.66

414726	BE466863	Hs.280099	ESTs	9.66
418485	R91679	Hs.124981	ESTs	9.66
433480	X02422	Hs.181125	immunoglobulin lambda locus	9.65
441530	AI248301	Hs.127112	ESTs	9.65
433533	D53304	Hs.65394	ESTs	9.65
421470	R27496	Hs.1378	annexin A3	9.64
438613	C05569	Hs.243122	hypothetical protein FLJ13057 similar to	9.64
429324	AA488101	Hs.199245	inactivation escape 1	9.62
450244	AA007534	Hs.125062	ESTs	9.62
407660	AW063190	Hs.279101	ESTs	9.61
406554				9.60
426404	AA377607	Hs.273138	ESTs	9.58
447045	AW392394	Hs.278569	KIAA0064 gene product	9.58
449894	AK001578	Hs.24129	hypothetical protein FLJ10716	9.58
448376	AI494332	Hs.196963	ESTs	9.58
407902	AL117474	Hs.41181	Homo sapiens mRNA; cDNA DKFZp727C191 (fr	9.56
446572	AV659151	Hs.282961	ESTs	9.56
459245	BE242623	Hs.31939	manic fringe (Drosophila) homolog	9.55
423545	AP000692	Hs.129781	chromosome 21 open reading frame 5	9.54
414697	BE266134	Hs.76927	translocase of outer mitochondrial membr	9.54
410846	AW807057		gb:MR4-ST0062-031199-018-b03 ST0062 Homo	9.52
421181	NM_005574	Hs.184585	LIM domain only 2 (rhombotin-like 1)	9.52
427308	D26067	Hs.174905	KIAA0033 protein	9.52
415995	NM_004573	Hs.994	phospholipase C, beta 2	9.51
434846	AW295389	Hs.119768	ESTs	9.51
414342	AA742181	Hs.75912	Homo sapiens cDNA: FLJ22199 fis, clone H	9.50
416959	D28459	Hs.80612	ubiquitin-conjugating enzyme E2A (RAD6 h	9.50
443123	AA094538	Hs.6588	ESTs	9.50
439312	AA833902	Hs.270745	ESTs	9.48
449375	R07114	Hs.271224	ESTs	9.48
436357	AJ132085		gb:Homo sapiens mRNA for axonemal dynein	9.44
458723	AW137726	Hs.244352	ESTs, Moderately similar to laminin alph	9.44
457526	AW450584	Hs.192131	ESTs, Weakly similar to RIBB [H.sapiens]	9.43
404741				9.43
422409	NM_005428	Hs.116237	vav 1 oncogene	9.43
403708				9.42
408806	AW847814	Hs.289005	Homo sapiens cDNA: FLJ21532 fis, clone C	9.42
417380	T06809		gb:EST04698 Fetal brain, Stratagene (cat	9.42
422501	AA354690	Hs.144967	ESTs	9.42
426197	AA004410	Hs.167835	acyl-Coenzyme A oxidase 1, palmitoyl	9.42
452624	AU076606	Hs.30054	coagulation factor V (proaccelerin, labi	9.42
412110	AW893569		gb:RC0-NN0021-040400-021-c10 NN0021 Homo	9.41
414158	AA361623	Hs.288775	Homo sapiens cDNA FLJ13900 fis, clone TH	9.41
408101	AW968504	Hs.123073	CDC2-related protein kinase 7	9.40
414171	AA360328	Hs.865	RAP1A, member of RAS oncogene family	9.40
415947	U04045	Hs.78934	mutS (E. coli) homolog 2 (colon cancer,	9.40
426959	BE262745		gb:601153869F1 NIH_MGC_19 Homo sapiens c	9.39
417519	AI689987	Hs.177669	ESTs, Weakly similar to RMS1_HUMAN REGUL	9.39
457181	BE514362	Hs.296422	FK506-binding protein 3 (25kD)	9.39
402835				9.38
404632				9.38
446566	H95741	Hs.17914	Homo sapiens cDNA: FLJ22801 fis, clone K	9.37
455369	AW903533		gb:CM1-NN1031-060400-178-d05 NN1031 Homo	9.37
444001	AI095087	Hs.152299	ESTs, Moderately similar to ALU5_HUMAN A	9.36
458191	AI420611	Hs.127832	ESTs	9.36
431374	BE258532	Hs.251871	CTP synthase	9.34
429327	AA283981	Hs.199248	prostaglandin E receptor 4 (subtype EP4)	9.33
407061	X97748		gb:H.sapiens PTX3 gene promotor region.	9.33
416967	BE616731	Hs.80645	interferon regulatory factor 1	9.33
423013	AW875443	Hs.22209	secreted modular calcium-binding protein	9.33
439461	AA693960	Hs.103158	ESTs	9.33
418830	BE513731	Hs.88959	Human DNA sequence from clone 967N21 on	9.32
422763	AA033699	Hs.83938	ESTs, Moderately similar to MASP-2 [H.sa	9.32
442739	NM_007274	Hs.8679	cytosolic acyl coenzyme A thioester hydr	9.32
452859	AI300555	Hs.288158	Homo sapiens cDNA: FLJ23591 fis, clone L	9.32
403237				9.32
415000	AW025529	Hs.239812	ESTs, Weakly similar to CALM_HUMAN CALMO	9.31
417951	AW976410	Hs.289069	Homo sapiens cDNA: FLJ21016 fis, clone C	9.30
419066	Z98492	Hs.6975	PRO1073 protein	9.30

448443	AW167128	Hs.231934	ESTs	9.30
405125				9.30
409768	AW499566		gb:UI-HF-BR0p-aji-h-03-0-UI.r1 NIH_MGC_5	9.28
453708	AI191811	Hs.54629	ESTs	9.28
442271	AF000652	Hs.8180	syndecan binding protein (syntenin)	9.27
410055	AJ250839	Hs.58241	gene for serine/threonine protein kinase	9.26
448692	AW013907	Hs.224276	ESTs, Moderately similar to predicted us	9.26
417381	AF164142	Hs.82042	solute carrier family 23 (nucleobase tra	9.25
422497	D29642	Hs.1528	KIAA0053 gene product	9.25
414140	AA281279	Hs.23317	ESTs	9.24
435980	AF274571	Hs.129142	ESTs; Weakly similar to DEOXYRIBONUCLEAS	9.24
458530	BE395035	Hs.199869	ESTs, Weakly similar to KIAA0874 protein	9.24
402585				9.24
420819	AA280700		gb:zs95h11.s1 NCI_CGAP_GCB1 Homo sapiens	9.23
444755	AA431791	Hs.183001	ESTs	9.22
411630	U42349	Hs.71119	Putative prostate cancer tumor suppresso	9.22
421246	AW582962	Hs.300961	ESTs, Highly similar to AF151805 1 CGI-4	9.20
421924	BE514514	Hs.109606	coronin, actin-binding protein, 1A	9.19
414888	AL039185	Hs.77558	thyroid hormone receptor interactor 7	9.18
434267	AI206589	Hs.116243	ESTs	9.17
409213	U61412	Hs.51133	PTK6 protein tyrosine kinase 6	9.17
428242	H55709	Hs.2250	leukemia inhibitory factor (cholinergic	9.16
451736	AW080356	Hs.293684	ESTs, Weakly similar to alternatively sp	9.15
413627	BE182082	Hs.246973	ESTs	9.14
416134	AA528402	Hs.74861	activated RNA polymerase II transcriptio	9.14
449251	AW151660	Hs.31444	ESTs	9.14
452813	U54727	Hs.191445	ESTs	9.14
443622	AI911527	Hs.11805	ESTs	9.14
413260	BE075281		gb:PM1-BT0585-290200-005-d07 BT0585 Homo	9.12
413450	Z99716	Hs.75372	N-acetylgalactosaminidase, alpha-	9.12
446442	BE221533	Hs.257858	ESTs	9.12
438540	AA810021	Hs.136906	ESTs	9.12
426251	M24283	Hs.168383	Intercellular adhesion molecule 1 (CD54)	9.11
410290	AA402307	Hs.73818	ubiquinol-cytochrome c reductase hinge p	9.10
437398	AA913736	Hs.126715	ESTs	9.10
421559	NM_014720	Hs.105751	Ste20-related serine/threonine kinase	9.10
439699	AF086534	Hs.187561	ESTs, Moderately similar to ALU1_HUMAN A	9.10
430799	C19035	Hs.164259	ESTs	9.09
424544	M88700	Hs.150403	dopa decarboxylase (aromatic L-amino aci	9.08
453942	AW190920	Hs.19928	ESTs	9.08
425844	T68073	Hs.159628	serine (or cysteine) proteinase inhibito	9.08
434658	AI624436	Hs.194488	ESTs	9.07
453999	BE328153	Hs.240087	ESTs	9.06
436490	R71543	Hs.18713	ESTs	9.05
409192	AA065131	Hs.233439	ESTs, Weakly similar to ALU7_HUMAN ALU S	9.05
446223	BE300091	Hs.119699	hypothetical protein FLJ12969	9.04
447247	AW369351	Hs.287955	Homo sapiens cDNA FLJ13090 fis, clone NT	9.04
450094	AI174947	Hs.295789	Homo sapiens mRNA; cDNA DKFZp564D1164 (f	9.04
432012	AW301344	Hs.195969	ESTs	9.04
422520	AU076730	Hs.117977	kinesin 2 (60-70kD)	9.02
418650	BE386750	Hs.86978	prolyl endopeptidase	9.02
423008	M81590	Hs.123016	5-hydroxytryptamine (serotonin) receptor	9.02
436476	AA326108	Hs.53631	ESTs	9.02
448206	BE622585	Hs.3731	ESTs	9.02
431574	AW572659	Hs.261373	adenosine A2b receptor pseudogene	9.01
443453	R99876	Hs.269882	ESTs	9.01
435472	AW972330	Hs.283022	triggering receptor expressed on myeloid	9.01
420337	AW295840	Hs.14555	Homo sapiens cDNA: FLJ21513 fis, clone C	9.00
449810	AB008681	Hs.23994	activin A receptor, type IIB	9.00
406780	AA902366	Hs.286	ribosomal protein L4	8.99
429169	AW341130	Hs.197757	ESTs, Moderately similar to FGFE_HUMAN F	8.99
421326	AF051428	Hs.103504	estrogen receptor 2 (ER beta)	8.97
425491	AA883316	Hs.255221	ESTs	8.96
425516	BE000707	Hs.29567	ESTs	8.96
439773	AI051313	Hs.143315	ESTs	8.96
443247	BE614387	Hs.47378	ESTs	8.96
456623	AI084125	Hs.108106	transcription factor	8.95
438707	L08239	Hs.5326	porcupine	8.95
402240				8.95

444152	AI125694	Hs.149305	Homo sapiens cDNA FLJ14264 fis, clone PL	8.95
409842	AW501756		gb:UI-HF-BR0p-ajm-c-09-0-UI.r1 NIH_MGC_5	8.94
416277	W78765	Hs.73580	ESTs	8.94
456697	AI908006	Hs.111334	ferritin, light polypeptide	8.94
410762	AF226053	Hs.66170	HSKM-B protein	8.92
412942	AL120344	Hs.75074	mitogen-activated protein kinase-activat	8.92
442320	AI287617	Hs.129636	ESTs	8.92
449673	AA002064	Hs.18920	ESTs	8.91
411486	N85785	Hs.181165	eukaryotic translation elongation factor	8.90
437916	BE566249	Hs.20999	Homo sapiens cDNA: FLJ23142 fis, clone L	8.90
442732	AA257161	Hs.8658	hypothetical protein DKFZp434E0321	8.89
419741	NM_007019	Hs.93002	ubiquitin carrier protein E2-C	8.89
411499	AW849292		gb:IL3-CT0215-020300-090-E06 CT0215 Homo	8.89
431154	AW971228	Hs.290259	ESTs	8.89
414922	D00723	Hs.77631	glycine cleavage system protein H (amino	8.88
418036	Z37976	Hs.83337	latent transforming growth factor beta b	8.87
406422				8.87
422926	NM_016102	Hs.121748	ring finger protein 16	8.87
435220	D50030	Hs.104	HGF activator	8.86
418203	X54942	Hs.83758	CDC28 protein kinase 2	8.86
418613	AA744529	Hs.86575	mitogen-activated protein kinase kinase	8.85
439250	H66566	Hs.271711	ESTs	8.85
432359	AA076049	Hs.274415	Homo sapiens cDNA FLJ10229 fis, clone HE	8.84
450000	AI952797	Hs.10888	Homo sapiens cDNA: FLJ21559 fis, clone C	8.83
425657	T89839	Hs.119471	ESTs	8.83
425694	U51333	Hs.159237	hexokinase 3 (white cell)	8.82
419972	AL041465	Hs.294038	ESTs, Moderately similar to ALU2_HUMAN A	8.82
436396	AI683487	Hs.299112	Homo sapiens cDNA FLJ11441 fis, clone HE	8.82
413413	D82520	Hs.301834	Homo sapiens cDNA FLJ10952 fis, clone PL	8.82
428807	AA435997	Hs.104930	ESTs	8.82
415839	R40611	Hs.137565	ESTs	8.81
419553	N34145	Hs.250614	ESTs	8.80
420309	AW043637	Hs.21766	ESTs	8.80
421863	AI952677	Hs.108972	Homo sapiens mRNA; cDNA DKFZp434P228 (fr	8.80
447965	AW292577	Hs.94445	ESTs	8.80
459172	BE063380		gb:PM0-BT0275-291099-002-g10 BT0275 Homo	8.80
403259				8.78
411534	AW850473		gb:IL3-CT0219-280100-061-B11 CT0219 Homo	8.78
456161	BE264645	Hs.282093	Homo sapiens cDNA: FLJ21918 fis, clone H	8.77
413654	AA331881	Hs.75454	peroxiredoxin 3	8.76
401744				8.76
425348	AL137477	Hs.155912	cadherin-like 24	8.76
423396	AI382555	Hs.127950	bromodomain-containing 1	8.75
450649	NM_001429	Hs.297722	Human DNA sequence from clone RP1-85F18	8.75
408331	NM_007240	Hs.44229	dual specificity phosphatase 12	8.74
423872	AB020316	Hs.134015	uronyl 2-sulfotransferase	8.74
424906	AI566086	Hs.153716	Homo sapiens mRNA for Hmob33 protein, 3'	8.74
427596	AA449506	Hs.179765	Homo sapiens mRNA; cDNA DKFZp586H1921 (f	8.73
432488	AA551010	Hs.216640	ESTs	8.72
448980	AL137527	Hs.22703	Homo sapiens mRNA; cDNA DKFZp434P1018 (f	8.72
429455	AI472111	Hs.292507	ESTs	8.71
429855	AW385597	Hs.138902	ESTs, Weakly similar to B34087 hypotheti	8.71
441746	H59955	Hs.127829	ESTs	8.70
411945	AL033527	Hs.92137	v-myc avian myelocytomatosis viral oncog	8.70
413492	D87470	Hs.75400	KIAA0280 protein	8.70
435706	W31254	Hs.7045	GL004 protein	8.70
433741	AA609019	Hs.159343	ESTs	8.70
426340	Z97989	Hs.169370	FYN oncogene related to SRC, FGR, YES	8.69
422779	AA317036	Hs.41989	ESTs	8.67
449785	AI225235	Hs.288300	Homo sapiens cDNA: FLJ23231 fis, clone C	8.67
420144	AA811813	Hs.119421	ESTs	8.66
420235	AA256756	Hs.31178	ESTs	8.66
432606	NM_002104	Hs.3066	granzyme K (serine protease, granzyme 3;	8.66
425762	BE244076	Hs.159578	Homo sapiens mRNA for FLJ00020 protein,	8.65
427448	BE246449	Hs.2157	Wiskott-Aldrich syndrome (eczema-thrombo	8.64
418033	W68180	Hs.259855	Homo sapiens cDNA FLJ12507 fis, clone NT	8.64
429084	AJ001443	Hs.195614	splicing factor 3b, subunit 3, 130kD	8.64
417094	NM_006895	Hs.81182	histamine N-methyltransferase	8.64
457277	NM_004736	Hs.227656	xenotropic and polytropic retrovirus rec	8.63

422631	BE218919	Hs.118793	hypothetical protein FLJ10688	8.63
410679	AW795196	Hs.215857	ring finger protein 14	8.63
431585	BE242803	Hs.262823	hypothetical protein FLJ10326	8.62
401851				8.62
401866				8.62
407783	AW996872	Hs.172028	a disintegrin and metalloproteinase doma	8.62
408242	AA251594	Hs.43913	PIBF1 gene product	8.62
422250	AW408530	Hs.113823	ClpX (caseinolytic protease X, E. coli)	8.62
430259	BE550182	Hs.127826	RalGEF-like protein 3, mouse homolog	8.62
452598	AI831594	Hs.68647	ESTs, Weakly similar to ALU7_HUMAN ALU S	8.62
419541	AW749617		gb:RC3-BT0502-130100-012-g07 BT0502 Homo	8.60
428839	AI767756	Hs.82302	ESTs	8.60
429328	AA829402	Hs.47939	ESTs	8.60
451491	AI972094	Hs.286221	Homo sapiens cDNA FLJ13741 fis, clone PL	8.60
452561	AI692181	Hs.49169	KIAA1634 protein	8.60
420027	AF009746	Hs.94395	ATP-binding cassette, sub-family D (ALD)	8.60
435205	X54136	Hs.181125	immunoglobulin lambda locus	8.60
430900	U91939	Hs.248123	G protein-coupled receptor 25	8.60
405074				8.59
437991	AI479773	Hs.181679	ESTs	8.59
436346	BE328882	Hs.193096	ESTs, Moderately similar to U119_HUMAN U	8.58
411079	AA091228		gb:cchn2152.seq.F Human fetal heart, Lam	8.57
418452	BE379749	Hs.85201	C-type (calcium dependent, carbohydrate-	8.56
429109	AL008637	Hs.196352	neutrophil cytosolic factor 4 (40kD)	8.56
448019	AW947164	Hs.195641	ESTs	8.56
449865	AW204272	Hs.199371	ESTs	8.55
431180	H55883		gb:yq94h03.r1 Soares fetal liver spleen	8.54
445988	BE007663	Hs.13503	inactivation escape 2	8.54
405876				8.54
407235	D20569	Hs.169407	SAC2 (suppressor of actin mutations 2, y	8.54
414807	AI738616	Hs.77348	hydroxyprostaglandin dehydrogenase 15-(N	8.54
425671	AF193612	Hs.159142	lunatic fringe (Drosophila) homolog	8.54
452413	AW082633	Hs.212715	ESTs	8.54
421620	AA446183	Hs.91885	ESTs	8.53
444539	AI955765	Hs.146907	ESTs	8.52
415102	M31899	Hs.77929	excision repair cross-complementing rode	8.51
405552				8.51
418068	AW971155	Hs.293902	ESTs, Weakly similar to prolyl 4-hydroxy	8.50
420133	AA426117	Hs.14373	ESTs	8.50
438887	R68857	Hs.265499	ESTs	8.50
446468	AI765890	Hs.16341	ESTs; Moderately similar to !!!! ALU SUB	8.50
446585	AV659397	Hs.282948	ESTs	8.50
441896	AW891873		gb:CM3-NT0090-040500-173-b02 NT0090 Homo	8.50
437718	AI927288	Hs.196779	ESTs	8.48
420656	AA279098	Hs.187636	ESTs	8.48
429303	AW137635	Hs.44238	ESTs	8.48
450624	AL043983	Hs.125063	Homo sapiens cDNA FLJ13825 fis, clone TH	8.48
452573	AI907957	Hs.287622	Homo sapiens cDNA FLJ14082 fis, clone HE	8.48
456341	AA229126	Hs.122647	N-myristoyltransferase 2	8.48
423024	AA593731	Hs.75613	CD36 antigen (collagen type I receptor,	8.47
446985	AL038704	Hs.156827	ESTs, Weakly similar to ALU1_HUMAN ALU S	8.46
431778	AL080276	Hs.268562	regulator of G-protein signalling 17	8.46
400268				8.46
421828	AW891965	Hs.289109	dimethylarginine dimethylaminohydrolase	8.45
417022	NM_014737	Hs.80905	Ras association (RalGDS/AF-6) domain fam	8.44
421029	AW057782	Hs.293053	ESTs	8.44
425171	AW732240	Hs.300615	ESTs	8.44
459070	AI814302		gb:wj71c12.x1 NCI_CGAP_Lu19 Homo sapiens	8.42
406006				8.42
412643	AW971239	Hs.293982	ESTs	8.42
424775	AB014540	Hs.153026	SWAP-70 protein	8.42
446848	AW136083	Hs.195266	ESTs, Weakly similar to S59501 interfero	8.42
448043	AI458653	Hs.201881	ESTs	8.41
407183	AA358015		gb:EST66864 Fetal lung III Homo sapiens	8.40
412324	AW978439	Hs.69504	ESTs	8.40
419594	AA013051	Hs.91417	topoisomerase (DNA) II binding protein	8.40
430968	AW972630		gb:EST384925 MAGE resequences, MAGL Homo	8.40
431689	AA305688	Hs.267695	UDP-Gal:betaGlcNAc beta 1,3-galactosyltr	8.40
438582	AI521310	Hs.283365	ESTs, Weakly similar to ALU5_HUMAN ALU S	8.40

	447685	AL122043	Hs.19221	hypothetical protein DKFZp566G1424	8.40
	459119	AW844498	Hs.289052	Homo sapiens LENG8 mRNA, variant C, part	8.38
	400817				8.37
	425265	BE245297		gb:TCBAP1E2482 Pediatric pre-B cell acut	8.37
5	409385	AA071267		gb:zm61g01.r1 Stratagene fibroblast (937	8.36
	439121	BE047779	Hs.44701	ESTs	8.36
	419968	X04430	Hs.93913	interleukin 6 (interferon, beta 2)	8.36
	408327	AW182309	Hs.249963	ESTs, Highly similar to dJ1170K4.4 [H.sa	8.35
	403976				8.34
10	448064	AA379036		gb:EST91809 Synovial sarcoma Homo sapien	8.33
	442914	AW188551	Hs.99519	Homo sapiens cDNA FLJ14007 fis, clone Y7	8.33
	428032	AW997704	Hs.11493	Homo sapiens cDNA FLJ13536 fis, clone PL	8.32
	434194	AF119847	Hs.283940	Homo sapiens PRO1550 mRNA, partial cds	8.32
	458677	AW937670	Hs.254379	ESTs	8.32
15	420925	NM_015698	Hs.100391	T54 protein	8.30
	416475	T70298		gb:yd26g02.s1 Soares fetal liver spleen	8.30
	416852	AF283776	Hs.80285	Homo sapiens mRNA; cDNA DKFZp566C1723 (f	8.30
	430676	AF084866		gb:Homo sapiens envelope protein RIC-3 (	8.30
	428455	AI732694	Hs.98520	ESTs	8.29
20	435343	AW194962	Hs.199028	ESTs	8.29
	450783	BE266695		gb:601190242F1 NIH_MGC_7 Homo sapiens cD	8.29
	404946				8.28
	422942	AF054839	Hs.122540	tetraspan 2	8.28
	453716	AA037675	Hs.152675	ESTs	8.28
25	437098	AA744488	Hs.132842	ESTs, Moderately similar to ALU1_HUMAN A	8.28
	443907	AU076484	Hs.9963	TYRO protein tyrosine kinase binding pro	8.27
	401930	AF106069	Hs.23168	ubiquitin specific protease 15	8.26
	446554	AA151730	Hs.301789	ESTs, Weakly similar to similar to C.ele	8.26
	426290	AB007918	Hs.169182	KIAA0449 protein	8.25
30	419904	AA974411	Hs.18672	ESTs	8.25
	413886	AW958264	Hs.103832	ESTs, Weakly similar to TRHY_HUMAN TRICH	8.24
	424738	AI963740	Hs.46826	ESTs	8.24
	427359	AW020782	Hs.79881	Homo sapiens cDNA: FLJ23006 fis, clone L	8.24
	424534	D87682	Hs.150275	KIAA0241 protein	8.24
35	424429	U63830	Hs.146847	TRAF family member-associated NFKB activ	8.24
	442604	BE263710	Hs.279904	ESTs	8.22
	442992	AI914699	Hs.13297	ESTs	8.22
	427210	BE396283	Hs.173987	eukaryotic translation initiation factor	8.22
	457229	BE222450	Hs.266390	ESTs	8.21
40	423730	AA330214		gb:EST33935 Embryo, 12 week II Homo sapi	8.21
	411928	AA888624	Hs.19121	adaptor-related protein complex 2, alpha	8.20
	416051	AA835868	Hs.25253	Homo sapiens cDNA: FLJ20935 fis, clone A	8.20
	417231	R40739	Hs.21326	ESTs	8.20
	422049	W25760	Hs.77631	glycine cleavage system protein H (amino	8.20
45	427528	AU077143	Hs.179565	minichromosome maintenance deficient (S.	8.20
	458776	AV654978	Hs.19904	cystathionase (cystathionine gamma-lyase	8.19
	417687	AI828596	Hs.250691	ESTs	8.18
	423218	NM_015896	Hs.167380	BLu protein	8.18
	425397	J04088	Hs.156346	topoisomerase (DNA) II alpha (170kD)	8.18
50	406964	M21305	Hs.247946	Human alpha satellite and satellite 3 ju	8.18
	402401	U42349	Hs.71119	Putative prostate cancer tumor suppresso	8.18
	423397	NM_001838	Hs.1652	chemokine (C-C motif) receptor 7	8.18
	427857	AL133017	Hs.2210	thyroid hormone receptor interactor 3	8.17
	401519				8.17
55	447188	H65423	Hs.17631	Homo sapiens cDNA FLJ20118 fis, clone CO	8.16
	424704	AI263293	Hs.152096	cytochrome P450, subfamily IIJ (arachido	8.16
	435854	AJ278120	Hs.4996	DKFZP564D166 protein	8.14
	448556	AW885606	Hs.5064	ESTs	8.14
	449217	AA278536	Hs.23262	ribonuclease, RNase A family, k6	8.14
60	453124	AI139058	Hs.23296	ESTs	8.14
	442812	AI018406	Hs.131284	ESTs	8.14
	421129	BE439899	Hs.89271	ESTs	8.14

TABLE 9A shows the accession numbers for those primekeys lacking a unigeneID in Table 9. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Pkey:	Unique Eos probeset identifier number
CAT number:	Gene cluster number
Accession:	Genbank accession numbers

Pkey	CAT number	Accession
408057	1035720_-1	AW139565
408069	103655_1	H81795 Z42291 R20973 AA046920
408182	104479_1	AA047854 AA057506 AA053841
408338	1052148_1	AW867079 AW867086 AW182772
408828	108463_1	BE540279 AW410659 AA057857 R77693 BE278674
409126	110159_1	AA063426 AW962323 AW408063 AA063503 AA772927 AW753492 BE175371 AA311147
409292	111586_1	AA071051 AA070584 AA069938 AA102136 AA074430
409314	111841_1	AA070266 AA084967 AA126998
409385	112523_1	AA071267 T65940 T64515 AA071334
409398	1126716_1	AW386461 AW876408 AW386672 AW386599 AW876258 AW386619 AW386289 AW876136 AW876203 AW876213 AW876301 AW876295 AW876349 AW876365 AW876160 AW876369 AW876352 AW876271
409671	114731_1	AA076769 AA076781 AI087968
409768	1154035_1	AW499566 AW502378 AW499522 AW502046 AW502671 AW501917 AW501868 AW501721 AW502813
409841	1156088_1	AW502139 AW502432 AW502235 AW501683 AW502647
409842	1156119_1	AW501756 AW502096 AW502465 AW501715
409853	1156226_1	AW502327 AW502488 AW501829 AW502625 AW502687
410531	1207200_1	AW752953 H88044 BE156092
410688	1216101_1	AW796342 AW796356 BE161430
410846	1223902_1	AW807057 AW807054 AW807189 AW807193 AW807369 AW807429 AW807364 AW807365 AW807078 AW807256 AW807180 AW807331
410896	1226053_1	AW809637 AW809697 AW810554 AW809707 AW809985 AW810000 AW810088 AW809742 AW809816 AW809749 AW809639 AW809722 AW809836 AW809774 AW810023 AW810013 AW809813 AW809660 AW809728 AW809768 AW809951 AW809657 AW809954
411079	123128_1	AA091228 H71860 H71073
411424	1245497_1	AW845985 AW845991 AW845962
411499	1248105_1	AW849292 AW849431 AW849422 AW849428 AW849420 AW849424 AW849427
411507	1248607_1	AW850140 AW850195 AW850192
411534	1248827_1	AW850473 AW850471 AW850431 AW850523
411972	1268491_1	BE074959 AW880160
412110	1277844_1	AW893569 AW893571 AW893588 AW893593
412226	1284289_1	W26786 AW998612 AW902272
412257	1285376_1	AW903830 BE071916
412405	1293012_1	AW948126 AW948139 AW948196 AW948145 AW948162 AW948134 AW948127 AW948124 AW948153 AW948157 AW948125 AW948131 AW948158 AW948164 AW948151
413260	1356003_1	BE075281 BE075219 BE075123 BE075119 BE075046
413471	1371778_1	BE142098 BE142092
413729	1385114_1	BE159999 BE160056 BE160107 BE160139
414182	142409_1	AA136301 AI381776 AA136321
414989	1511339_1	T81668 C19040 C17569
415354	1534763_1	F06495 R24336 R13046
416011	1566439_1	H14487 R50911 Z43216
416475	1596398_1	T70298 H58072 R02750
417380	1672461_1	T06809 N75735
419392	1843934_-1	W28573
419541	185724_1	AW749617 R64714 AA244138 AA244137 BE094019
419544	185760_2	AI909154 AA526337 AA244193 AI909153
420819	196721_1	AA280700 AW975494 AA687385
421245	200620_1	AA285363 AA285333 AA285359 AA285326 AA285350
422673	219674_1	N59027 AA314694 N53937 R08100

422695	219996_1	AA315158 AW961298 N76067 AW802759 AI858495 W04474
422858	222209_1	R35398 BE252178 AA318153
422940	223106_1	BE077458 AA337277 AA319285
423730	231462_1	AA330214 AW962519 T54709
423790	232031_1	BE152393 AA330984 BE073904
424385	238731_1	AA339666 AW952809 AA349119
424606	241409_1	AA343936 AA344060 AW963081
425265	249175_1	BE245297 AA353976 AW505023
426959	273830_-1	BE262745
430676	32168_1	AF084866 AF084870 AF084864 AF084867 AF084869 AF084865 AF084868 AW818206 AW812038 BE144813 BE144812 AW812041 AW812040 AW812067 BE061583 BE061604 T05808 AI352469 AA580921 BE141783 BE141782 BE061601 AW814393 AW885029
430968	326269_1	AW972830 AA527647 AA489820 AA570362
431180	328906_1	H55883 AW971249 AA493900 H55786
432093	341283_1	H28383 AW972670 H28359 AA525808
434596	36937_1	T59538 T59589 T59598 T59542 AF147374
436357	41842_1	AJ132085 Z83805
437159	43393_1	AL050072 AW900148
437495	43765_1	BE177778 BE177779 AL390180 AA359908
439097	46858_1	H66948 AF085954 H66949
439120	46879_1	H56389 AF085977 H56173
440134	48675_1	BE410734 BE560117 BE270054 BE296330 BE267957 AI003007 BE545259
441896	52842_1	AW891873 AW891897 BE564764
445629	645767_1	AI245701 BE272724
447229	71288_1	BE617135 AW504051 AW504283
448064	74761_1	AA379036 AA150589 AI696854 BE621316
450783	84655_1	BE266695 BE265474 N53200 BE267333
451045	85673_1	AA215672 AI696628 AA013335 H86334 AA017006
452549	921802_1	AI907039 AI907081
452560	922216_1	BE077084 AW139963 AW863127 AW806209 AW806204 AW806205 AW806206 AW806211 AW806212 AW806207 AW806208 AW806210 AI907497
452712	928309_1	AW838616 AW838660 BE144343 AI914520 AW888910 BE184854 BE184784
453758	980026_1	U83527 AL120938 U83522
454093	1007366_1	AW860158 AW862385 AW860159 AW862386 AW862341 AW821869 AW821893 AW062660 AW062656
454563	1224342_1	AW807530 AW807540 AW807537 AW846086 BE141634 AW846089 AW807499 AW807533 AW838499
454791	1234759_1	BE071874 BE071882 AW820782 AW821007
454977	1247099_1	AW848032 AW848630 AW848478 AW848623 AW848484 AW848169 AW848830 AW848149 AW848119 AW848893 AW848903 AW848407
455131	1254674_1	AW857913 AW857916 AW857914 AW861627 AW861626 AW861624
455183	1259023_1	AW984111 AW863918 AW863856
455254	1266449_1	AW877015 AW877133 AW876978 AW877071 AW876988 AW877069 AW877063 AW877013
455369	1285173_16	AW903533 AW903516 AW903562 BE085202 BE085215 BE085214 BE085209 BE085172 BE085175 BE085193 BE085211 BE085199
455982	1396849_1	BE176862 BE176876 BE176947 BE176878
456011	1410860_1	BE243628 BE246081 BE247016 BE241984 BE241534 BE246091 BE245679 BE243620 BE245998 BE242329 BE241417 BE241457 BE242522 BE241989 BE241464
456023	1416335_1	R00028 BE247630
457586	360505_1	AW062439 AW751554 AA579463
457595	364225_-1	AA584854
457751	399422_1	AI908236 AA663731
459070	883688_1	AI814302 AI814428
459081	889426_1	W07808 AI822066
459145	918957_1	AI903354 AI903489 AI903488
459172	921149_1	BE063380 BE063346 AI906097
459234	945240_-1	AI940425

TABLE 9B shows the genomic positioning for those primekeys lacking unigene ID's and accession numbers in Table 9. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

Pkey:	Unique number corresponding to an Eos probeset
Ref:	Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-495.
Strand:	Indicates DNA strand from which exons were predicted.
Nt_position:	Indicates nucleotide positions of predicted exons.

Pkey	Ref	Strand	Nt_position
400452	8113550	Minus	90308-90505
400557	9801261	Plus	208453-208528,209633-209813
400615	9908994	Plus	118036-118166,118681-118807
400802	8567867	Minus	174571-174856
400817	8569994	Plus	170793-170948
400880	9931121	Plus	29235-29336,36363-36580
400885	9958187	Minus	58242-58733
400926	7651921	Minus	52033-52158,53956-54120,54957-55052,55420-55480,56452-56666,57221-57718
400952	7658481	Plus	192667-192826,194387-194876
400991	8096825	Plus	159197-159320
401044	8117619	Plus	73501-73674
401124	8570296	Minus	124181-124391
401163	6981820	Plus	5302-5545
401201	9743387	Minus	138534-138629,139234-139294,140121-140335,142033-142479
401286	9801342	Minus	147036-147318
401384	6850939	Minus	58360-58545
401468	6433826	Plus	13056-13482
401515	7630851	Plus	29929-30126
401519	6649315	Plus	157315-157950
401672	9838136	Plus	128526-128704,130755-130860
401744	2576349	Plus	14595-14751
401851	7770425	Minus	146443-146664,147794-147971,148351-148480,148980-149111,149801-149949
401866	8018106	Plus	73126-73623
402240	7690131	Plus	104382-104527,106136-106372
402359	9211204	Minus	40403-41961
402585	9908890	Minus	174893-175050,183210-183435
402788	9796102	Plus	98273-101430
402802	3287156	Minus	53242-53432
402812	6010110	Plus	25026-25091,25844-25920
402828	8918414	Plus	69071-69642
402835	9187337	Plus	26961-27101
402838	9369121	Minus	32589-32735,35478-35666
402842	9369121	Minus	76355-76479
402895	9967547	Plus	85537-85671,86379-86469
402964	9581599	Minus	46624-46784
403137	9211494	Minus	92349-92572,92958-93084,93579-93712,93949-94072,94591-94748,95214-95337
403237	7637607	Plus	7271-7527
403259	7770585	Plus	4693-4857
403683	7331517	Plus	217175-217446
403690	7387384	Minus	78627-79583
403708	5705981	Minus	134394-134812
403838	4176355	Plus	19197-19502
403851	7709872	Plus	22733-23007
403976	7657840	Plus	24755-24969
404407	7329316	Minus	48154-48499
404426	7407959	Plus	77842-77954
404632	9796668	Plus	45096-45229
404741	8574139	Plus	143025-143467
404756	7706327	Plus	82849-83627
404946	7382189	Plus	134445-134750
405074	7770440	Plus	44340-44559,44790-45059
405125	8247873	Plus	137113-137814
405172	9966752	Plus	153027-153262

405236	7249076	Minus	151699-151915
405325	6094661	Minus	25818-26380
405411	3451356	Minus	17503-17778,18021-18290
405495	8050952	Minus	72182-72373
405552	1552506	Plus	45199-45647
405601	5815493	Minus	147835-147935,149220-149299
405685	4508129	Minus	37956-38097
405777	7263187	Minus	104773-105051
405856	7653009	Plus	101777-102043
405876	6758747	Plus	39694-40031
405932	7767812	Minus	123525-123713
405934	6758795	Plus	159913-160605
406006	8247801	Minus	42640-42776
406134	9163473	Plus	153291-153452
406189	7289992	Minus	22007-22234
406422	9256411	Plus	163003-163311
406516	7711422	Minus	128375-128449,128560-128784
406538	7711478	Plus	35196-35367,38229-38476,40080-40216,43522-43840
406554	7711566	Plus	106956-107121
406577	7711730	Plus	11377-11509

TABLE 10: shows genes, including expression sequence tags differentially expressed in taxol resistant prostate tumor xenografts as compared to taxol sensitive prostate tumor xenografts. The genes are indicated as either being upregulated or downregulated during the induction of taxol resistance in sequential passages of the grafts.

Pkey:	Unique Eos probeset identifier number
ExAccn:	Exemplar Accession number, Genbank accession number
UnigeneID:	Unigene number
Unigene Title:	Unigene gene title
Eos:	Internal Eos name
F00-F14:	passage number

Pkey	ExAccn	UnigeneID	UnigenTitle	Eos	Resp.F00	F00	F02	F02	F05	F05	F07	F09	F10	F11	F13	F14		
117921	N51002	Hs.47170	Liprin A2	PM28UP	1	9	8	9	32	20	34	122	105	82	71	111		
112971	T17185	Hs.4299	ESTs	CHA1down	290	281	267	335	270	284	150	157	83	89	49	75		
126645	AI167942	Hs.61635	STEAP	PAA5down	106	111	103	71	34	67	33	14	2	1	1	1		
119018	N95796	Hs.179809	ESTs	PAB2down	765	841	757	909	742	704	478	428	253	175	228	238		
110844	N31952	Hs.167531	ESTs	PAV7down	175	192	147	141	123	129	73	65	55	48	54	84		
100654	HG2841-HT2969		Hs.75442	Albumin, A	PM01down	666	605	504	728	357	445	602	187	117	127	117	113	
100655	HG2841-HT2970		Hs.75442	Albumin, A	PM02down	620	653	486	688	368	386	606	175	101	95	115	97	
102076	U09579	Hs.252437	cyclin-dep	PM03down	101	94	143	190	105	107	88	40	34	31	46	22		
102208	U22961	Hs.75442	albumin	PM04down	495	424	323	518	252	296	467	188	169	143	165	145		
103739	AA075779	--	mitochondr	PM05down	75	190	606	230	378	106	218	88	69	192	69	99		
107036	AA599690	Hs.15725	SBBI48	PM06down	87	124	115	188	132	111	66	71	49	70	38	50		
108242	AA062746	--	ESTs	PM07down	14	20	252	13	22	43	193	10	10	104	21	18		
108282	AA065143	--	solute car	PM08down	27	54	178	73	108	37	53	24	14	53	15	34		
108679	AA115963	--	beta-1-glo	PM09down	680	893	1292	656	669	389	1	74	118	662	359	409		
108731	AA126313	Hs.107476	ATP syntha	PM10down	10	19	185	25	60	1	32	3	7	14	1	1		
110675	H89355	Hs.6598	adrenergic	PM11down	207	334	237	239	231	220	119	145	93	64	56	124		
115412	AA283804	Hs.193552	ESTs	PM12down	146	316	282	271	340	334	115	238	100	196	83	207		
115844	AA430124	Hs.234607	MDM2	PM13down	49	93	94	154	132	91	23	54	23	76	14	41		
120588	AA281591	Hs.16193	ESTs	PM14down	80	157	58	141	159	127	39	83	35	37	16	46		
132349	Y00705	Hs.181286	serine pro	PM15down	146	217	214	150	106	128	177	85	54	63	66	56		
132868	AA490775	Hs.5920	N-acetylma	PM16down	92	150	132	178	126	139	53	94	48	67	41	80		
132967	AA032221	Hs.61635	STEAP	PM17down	224	208	203	215	205	180	132	65	68	50	48	63		
133063	AA283085	Hs.64065	ESTs	PM18down	85	148	161	150	92	108	42	99	42	65	29	126		
134374	D62633	Hs.8236	ESTs	PM19down	230	240	194	212	231	189	89	123	107	95	68	91		
135400	M23263	Hs.99915	androgen r	PM20down	36	167	99	178	132	101	23	71	26	122	14	44		

TABLE 11: shows genes, including expression sequence tags that are up-regulated in prostate tumor tissue compared to normal prostate tissue as analyzed using Affymetrix/Eos Hu01 GeneChip array. Shown are the ratios of "average" normal prostate to "average" prostate cancer tissues.

Pkey:	Unique Eos probeset identifier number
ExAccn:	Exemplar Accession number, Genbank accession number
UnigeneID:	Unigene number
Unigene Title:	Unigene gene title
R1:	Background subtracted normal prostate : prostate tumor tissue

Pkey	ExAccn	UnigeneID	Unigene Title	R1
101336	L49169	Hs.75678	FBJ murine osteosarcoma viral oncogene homolog B	0.012
130642	M63438	Hs.156110	Immunoglobulin kappa variable 1D-8	0.015
133512	X01677	Hs.195188	glyceraldehyde-3-phosphate dehydrogenase	0.017
133436	H44631	Hs.737	immediate early protein	0.017
129292	X13810	Hs.1101	POU domain; class 2; transcription factor 2	0.019
100610	HG2566-HT4792		Microtubule-Associated Protein Tau, Alt. Splice 3, Exon 8	0.02
133448	M34516	Hs.170116	immunoglobulin lambda-like polypeptide 3	0.021
125193	W67577	Hs.84298	CD74 antigen (invariant polypeptide of major histocompatibility complex; class II antigen-associated)	0.022
133456	T49257	Hs.183704	ubiquitin C	0.022
134546	AA459310	Hs.8518	Homo sapiens mRNA; cDNA DKFZp586L1722 (from clone DKFZp586L1722)	0.023
102131	U15085	Hs.1162	major histocompatibility complex; class II; DM beta	0.023
101375	M13560	Hs.84298	CD74 antigen (invariant polypeptide of major histocompatibility complex; class II antigen-associated)	0.023
100674	HG3033-HT3194		Spliceosomal Protein Sap 62	0.024
134365	R32377	Hs.82240	syntaxin 3A	0.027
132335	D60387	Hs.189885	ESTs	0.027
110303	H37901	Hs.32706	ESTs	0.028
131678	N59162	Hs.30542	ESTs	0.028
116599	D80046	Hs.250879	ESTs	0.029
133769	M17733	Hs.75968	thymosin; beta 4; X chromosome	0.029
107904	AA026648	Hs.61389	ESTs	0.03
129427	T60746	Hs.111334	ferritin; light polypeptide	0.03
105987	AA406631	Hs.110299	mitogen-activated protein kinase kinase 7	0.03
131466	F03233	Hs.27189	ESTs	0.032
102859	X00274	Hs.76807	Human HLA-DR alpha-chain mRNA	0.032
134626	S82198	Hs.8709	caldecrin (serum calcium decreasing factor; elastase IV)	0.032
134170	M63138	Hs.79572	cathepsin D (lysosomal aspartyl protease)	0.033
131713	X57809	Hs.181125	immunoglobulin lambda gene cluster	0.034
100748	HG3517-HT3711		Alpha-1-Antitrypsin, 5' End	0.034
118769	N74496		ESTs	0.034
111734	R25375	Hs.126916	ESTs	0.036
109221	AA192755	Hs.85840	ESTs; Weakly similar to stac [H.sapiens]	0.036
133846	AA480073	Hs.76719	U6 snRNA-associated Sm-like protein	0.036
135281	AA401575	Hs.97757	ESTs	0.037
119073	R32894	Hs.45514	v-ets avian erythroblastosis virus E26 oncogene related	0.037
100760	HG3576-HT3779		Major Histocompatibility Complex, Class II Beta W52	0.037
101426	M19483	Hs.25	ATP synthase; H+ transprtng; mitochndrl F1 complex; beta polypept	0.038
129568	AA428025	Hs.114360	transforming growth factor beta-stimulated protein TSC-22	0.038
130900	Z38468	Hs.21036	ESTs; Moderately similar to F25965_3 [H.sapiens]	0.039
133879	M13829	Hs.77183	v-raf murine sarcoma 3611 viral oncogene homolog 1	0.039
100627	HG2702-HT2798		Serine/Threonine Kinase (Gb:Z25424)	0.039
129424	M55593	Hs.111301	matrix metalloproteinase 2 (gelatinase A; 72kD gelatinase; 72kD type IV collagenase)	0.039
128652	AA621245	Hs.103147	ESTs; Weakly similar to similar to SP:YR40_BACSU [C.elegans]	0.039
129979	T72635	Hs.13956	ESTs	0.039
133468	X03068	Hs.73931	major histocompatibility complex; class II; DQ beta 1	0.04
102636	U67092		Human ataxia-telangiectasia locus protein (ATM) gene, exons 1a, 1b, 2, 3 and 4, partial cds	0.04
129536	M33493	Hs.184504	tryptase; alpha	0.04
133599	M64788	Hs.75151	RAP1; GTPase activating protein 1	0.041

	102104	U12139		Human alpha1(XI) collagen (COL11A1) gene, 5' region and exon 1	0.041
	131340	AA478305	Hs.25817	Homo sapiens chromosome 19; cosmid R27216	0.041
	130446	X79510	Hs.155693	protein tyrosine phosphatase; non-receptor type 21	0.042
	101352	L77701	Hs.16297	COX17 (yeast) homolog; cytochrome c oxidase assembly protein	0.042
5	122593	AA453310	Hs.128749	alpha-methylacyl-CoA racemase	0.042
	130181	R39552	Hs.151608	Homo sapiens clone 23622 mRNA sequence	0.042
	134071	Z14093	Hs.78950	branched chain keto acid dehydrogenase E1; alpha polypeptide (maple syrup urine disease)	0.042
	108129	AA053252	Hs.185848	ESTs; Weakly similar to !! ALU SUBFAMILY J WARNING	
10				ENTRY !! [H.sapiens]	0.043
	130511	L32137	Hs.1584	cartilage oligomeric matrix protein (pseudoachondroplasia; epiphyseal dysplasia 1; multiple)	0.043
	133336	AA291456	Hs.71190	ESTs	0.043
	132982	L02326	Hs.198118	immunoglobulin lambda-like polypeptide 2	0.044
15	131880	AA047034	Hs.33818	RecQ protein-like 5	0.044
	130540	U35234	Hs.159534	protein tyrosine phosphatase; receptor type; S	0.044
	133467	AA258595	Hs.73931	major histocompatibility complex; class II; DQ beta 1	0.044
	101191	L20688	Hs.83656	Rho GDP dissociation inhibitor (GDI) beta	0.044
	101860	M95610	Hs.37165	collagen; type IX; alpha 2	0.044
20	102799	U88898		Human endogenous retroviral H protease/integrase-derived ORF1 mRNA, complete cds, and putative envelope prot mRNA, partial cds	0.044
	107200	D20350	Hs.5628	ESTs	0.044
	101166	L14927	Hs.2099	lipocalin 1 (protein migrating faster than albumin; tear prealbumin)	0.044
	134289	M54915	Hs.81170	pim-1 oncogene	0.044
25	135329	AA436026	Hs.98858	ESTs	0.044
	124950	T03786	Hs.151531	protein phosphatase 3 (formerly 2B); catalytic subunit; beta isoform (calcineurin A beta)	0.044
	102919	X12447	Hs.183760	aldolase A; fructose-bisphosphate	0.044
	100574	HG2279-HT2375		Triosephosphate Isomerase	0.045
30	131286	AA450092	Hs.25300	Homo sapiens clones 24718 and 24825 mRNA sequence	0.045
	102675	U72512		Human B-cell receptor associated protein (hBAP) alternatively spliced mRNA, partial 3'UTR	0.045
	131332	R50487	Hs.25717	ESTs	0.045
	101634	M57731	Hs.75765	GRO2 oncogene	0.046
35	113118	T47906	Hs.220512	ESTs	0.046
	124884	R77276	Hs.120911	ESTs	0.046
	130523	W76097	Hs.214507	ESTs	0.046
	110244	H26742	Hs.25367	ESTs; Weakly similar to ALR [H.sapiens]	0.046
	131932	AA454980	Hs.25601	chromodomain helicase DNA binding protein 3	0.046
40	132509	H09751	Hs.5038	neuropathy target esterase	0.046
	133372	AA291139	Hs.72242	ESTs	0.046
	100817	HG4011-HT4804		Dystrophin-Associated Glycoprotein, 50 Kda, Alt. Splice 2	0.047
	106746	AA476436	Hs.7991	ESTs	0.047
	135401	L14813	Hs.169271	carboxyl ester lipase-like (bile salt-stimulated lipase-like)	0.047
45	130479	R44163	Hs.12457	Homo sapiens clone 23770 mRNA sequence	0.047
	102589	U62015	Hs.8667	cysteine-rich; angiogenic inducer; 61	0.047
	121521	AA412165	Hs.97358	EST	0.048
	135340	AA425137	Hs.99093	Homo sapiens chromosome 19; cosmid R28379	0.048
	132336	AA342422	Hs.45073	ESTs	0.048
50	115368	AA282133	Hs.88960	ESTs; Weakly similar to similar to collagen [C.elegans]	0.048
	101278	L38487	Hs.110849	estrogen-related receptor alpha	0.048
	103284	X80200	Hs.8375	TNF receptor-associated factor 4	0.048
	100564	HG2239-HT2324		Potassium Channel Protein (Gb:Z11585)	0.048
	133132	Z40883	Hs.65588	ESTs; Weakly similar to dJ393P12.2 [H.sapiens]	0.048
55	121811	AA424535	Hs.98416	ESTs	0.048
	129613	AA279481	Hs.238831	ESTs; Weakly similar to collagen alpha 1(XVIII) chain [M.musculus]	0.049
	132468	S79854	Hs.49322	deiodinase; iodothyronine; type III	0.049
	120111	W95841	Hs.136031	ESTs	0.049
	103668	Z83741	Hs.248174	H2A histone family; member M	0.049
60	130386	F10874	Hs.234249	mitogen-activated protein kinase 8 interacting protein 1	0.049
	104275	C02170	Hs.39387	ESTs; Weakly smlr to weak smlrity to ribosomal prot L14 [C.elegans]	0.049
	106305	AA436146	Hs.12828	ESTs	0.05
	116431	AA609878	Hs.55289	ESTs; Weakly smlr to 110 KD CELL MEMBRANE GLYCOPROTEIN [H.sapiens]	0.813
	120339	AA206465	Hs.256470	EST	0.05
65	114427	AA017063		ESTs; Highly similar to Mlz-1 protein [H.sapiens]	0.05
	118821	N79070	Hs.94789	ESTs	0.05
	118979	N93798	Hs.43666	protein tyrosine phosphatase type IVA; member 3	0.05
	107495	W78776	Hs.90375	ESTs	0.051
	120240	Z41732	Hs.66049	ESTs	0.051

	114331	Z41309	Hs.12400	ESTs	0.051
	130947	R40037	Hs.21506	ESTs	0.052
	129242	W81679	Hs.5174	ribosomal protein S17	0.052
	131413	AA482390	Hs.26510	ESTs; Modly smlr to vacuolar prot sorting homolog r-vps33b [R.norvegicus]	0.052
5	112304	R54798	Hs.26239	ESTs	0.052
	101416	M17254	Hs.45514	v-ets avian erythroblastosis virus E26 oncogene related	0.052
	131201	AA426304	Hs.24174	ESTs	0.052
	101054	K02405	Hs.73933	Human MHC class II HLA-DQ-beta mRNA (DR7 DQw2); complete cds	0.052
	101306	L41143	Hs.232069	T-cell leukemia translocation altered gene	0.053
10	129311	T55087		yb45c08.r1 Stratagene fetal spleen (#937205) Homo sapiens cDNA clone IMAGE:74126 5', mRNA sequence.	0.053
	129942	U95301	Hs.144442	phospholipase A2; group X	0.053
	119210	R93340	Hs.92995	ESTs	0.053
	101046	K01160		Accession not listed in Genbank	0.053
15	114086	Z38266	Hs.12770	Homo sapiens PAC clone DJ0777O23 from 7p14-p15	0.053
	110171	H19964	Hs.31709	ESTs	0.053
	101004	J04101	Hs.248109	v-ets avian erythroblastosis virus E26 oncogene homolog 1	0.053
	129715	N58479	Hs.12126	ESTs; Weakly similar to LR8 [H.sapiens]	0.053
	101581	M34996	Hs.198253	major histocompatibility complex; class II; DQ alpha 1	0.053
20	113285	T66830	Hs.182712	ESTs	0.053
	127537	AA569531	Hs.162859	ESTs	0.054
	100813	HG3995-HT4265		Cpg-Enriched Dna, Clone S19	0.054
	101841	M93107	Hs.76893	3-hydroxybutyrate dehydrogenase (heart; mitochondrial)	0.054
	135053	R77159	Hs.93678	ESTs	0.054
25	101419	M17886	Hs.177592	ribosomal protein; large; P1	0.054
	119724	W69468	Hs.47622	ESTs	0.055
	102673	U72509		Human alternatively spliced B8 (B7) mRNA, partial sequence	0.055
	129877	AA248589	Hs.13094	ESTs; Weakly similar to ORF YGR101w [S.cerevisiae]	0.055
	114788	AA156737	Hs.103904	EST	0.055
30	123812	AA620607	Hs.111591	ESTs	0.055
	117669	N39237	Hs.44977	ESTs	0.055
	123782	AA610111	Hs.162695	EST	0.055
	102395	U41767	Hs.92208	a disintegrin and metalloproteinase domain 15 (metargidin)	0.055
	133795	M12529	Hs.169401	*apolipoprotein E*	0.055
35	123193	AA489228	Hs.136956	ESTs	0.056
	132595	AA253369	Hs.155742	glyoxylate reductase/hydroxypyruvate reductase	0.056
	104161	AA456471	Hs.7724	KIAA0963 protein	0.056
	115330	AA281145	Hs.88827	ESTs	0.056
	112893	T08000	Hs.194684	bassoon (presynaptic cytomatrix protein)	0.056
40	133475	L29217	Hs.73987	CDC-like kinase 3	0.056
	128699	K03207	Hs.103972	proline-rich protein BstNI subfamily 4	0.056
	102940	X13956	Hs.24998	Hu 12S RNA induced by poly(rI); poly(rC) and Newcastle disease virus	0.056
	131299	AA431464	Hs.25426	ESTs; Weakly similar to unknown [H.sapiens]	0.057
	102495	U51240	Hs.79356	Lysosomal-associated multispanning membrane protein-5	0.057
45	129594	R70379	Hs.115396	Human germline IgD chain gene; C-region; C-delta-1 domain	0.057
	118593	N69020	Hs.207689	EST	0.057
	126702	U54602	Hs.2785	keratin 17	0.057
	124386	N27368	Hs.212414	sema domain; immunoglobulin domain (Ig); short basic domain; secreted; (semaphorin) 3E	0.057
50	130538	M20786	Hs.159509	alpha-2-plasmin inhibitor	0.057
	114299	Z40782	Hs.22920	similar to S68401 (cattle) glucose induced gene	0.057
	115604	AA400378	Hs.49391	ESTs	0.057
	106052	AA416947	Hs.6382	ESTs; Highly similar to KIAA0612 protein [H.sapiens]	0.057
	131730	U05681	Hs.31210	B-cell CLL/lymphoma 3	0.057
55	131285	AA479498	Hs.25274	ESTs; Modly smlr to putative seven pass transmembrane prot [H.sapiens]	0.058
	129705	X78706	Hs.12068	carnitine acetyltransferase	0.058
	123175	AA489010	Hs.178400	ESTs	0.058
	103592	Z30644	Hs.123059	chloride channel Kb	0.058
	118196	N59478	Hs.48396	ESTs; Moderately similar to tumor necrosis factor-alpha -induced protein B12 [H.sapiens]	0.058
60	104886	AA053348	Hs.144626	growth differentiation factor 11	0.058
	104250	AF000575	Hs.105928	leukocyte immunoglobulin-like receptor; subfamily B (with TM and ITIM domains); member 3	0.058
	113301	T67452	Hs.13104	EST	0.058
65	110441	H50302	Hs.19845	ESTs; Highly smlr to prot phosphatase 2A BR gamma subunit [H.sapiens]	0.058
	125297	Z39215	Hs.159409	ESTs	0.058
	135258	AA292423	Hs.97272	ESTs; Weakly similar to dJ281H8.2 [H.sapiens]	0.058
	130633	T92363	Hs.178703	ESTs	0.058
	112006	R42607	Hs.22241	hypothetical protein	0.058

	130805	U12194	Hs.170238	sodium channel; voltage-gated; type I; beta polypeptide	0.058
	134907	D80002	Hs.178292	KIAA0180 protein	0.058
	132619	AA404565	Hs.53447	ESTs; Moderately similar to kinesin light chain 1 [M.musculus]	0.058
	135115	N35489	Hs.94653	neurochondrin	0.058
5	100531	HG1872-HT1907		Major Histocompatibility Complex, Dg	0.058
	124530	N62256	Hs.102727	EST	0.058
	119960	W87533	Hs.32699	ESTs; Moderately similar to LIV-1 protein [H.sapiens]	0.058
	132793	AA478999	Hs.56966	KIAA0906 protein	0.058
	101076	L04270	Hs.1116	lymphotoxin beta receptor (TNFR superfamily; member 3	0.058
10	130655	N92934	Hs.17409	cysteine-rich protein 1 (intestinal)	0.058
	134458	AA192614	Hs.83577	cysteine and glycine-rich protein 3 (cardiac LIM protein)	0.058
	105904	AA401452	Hs.32060	ESTs	0.059
	132878	AA026793	Hs.58679	ESTs; Weakly similar to 4F2/CD98 light chain [M.musculus]	0.059
	121828	AA425166	Hs.98497	ESTs	0.059
15	133418	U76366	Hs.172727	Treacher Collins-Franceschetti syndrome 1	0.059
	129317	N46244	Hs.110373	ESTs	0.059
	130153	D85815	Hs.15114	ras homolog gene family; member D	0.059
	124403	N31745	Hs.102493	ESTs	0.059
	127683	AA668123	Hs.134170	ESTs	0.059
20	129814	W20070	Hs.168625	KIAA0979 protein	0.059
	131770	D59682	Hs.31833	ESTs	0.06
	117557	N33920	Hs.44532	diubiquitin	0.06
	103522	Y10514		H.sapiens mRNA for CD152 protein	0.06
	120029	W91960	Hs.250640	sequence-specific single-stranded-DNA-binding protein	0.06
25	102135	U15460	Hs.41691	activating transcription factor B	0.06
	123617	AA609183	Hs.181131	ESTs	0.06
	112136	R46100	Hs.9739	ESTs	0.061
	133725	V00563	Hs.179543	immunoglobulin mu	0.061
	102069	U09196	Hs.82520	Hu 1.1 kb mRNA upregltd in retinoic acid treated HL-60 neutrophilic cells	0.061
30	106555	AA455000	Hs.16725	ESTs	0.061
	123269	AA491226	Hs.105280	ESTs; Weakly similar to dJ963K23.2 [H.sapiens]	0.061
	109088	AA166837	Hs.72620	DKFZP434I114 protein	0.061
	129399	AA263028	Hs.111076	malate dehydrogenase 2; NAD (mitochondrial)	0.061
	129375	W79850	Hs.11081	ESTs; Weakly similar to HPBRII-7 protein [H.sapiens]	0.061
35	135271	AA397763	Hs.97562	ESTs	0.061
	132958	W90398	Hs.6147	KIAA1075 protein	0.061
	129364	AA477106	Hs.110757	DNA segment on chromosome 21 (unique) 2056 expressed sequence	0.061
	123427	AA598548	Hs.112471	ESTs	0.061
	105236	AA219179	Hs.19105	translocase of inner mitochondrial membrane 17 (yeast) homolog B	0.061
40	101012	J04444	Hs.697	cytochrome c-1	0.062
	134791	L18983	Hs.89655	protein tyrosine phosphatase; receptor type; N	0.062
	133700	K01396	Hs.75621	protease inhibitor 1 (anti-elastase); alpha-1-antitrypsin	0.062
	123887	AA621065	Hs.112943	ESTs	0.062
	129363	H05704	Hs.110746	H sapiens HCR (a-helix coiled-coil rod homologue) mRNA; complete cds	0.062
45	105719	AA291644	Hs.36793	ESTs	0.062
	124226	H62396	Hs.190266	ESTs	0.062
	117437	N27645		yw5e3.s1 Weizmann Olfactory Epithelium H sapiens cDNA clone IMAGE:255676 3' smlr to contains L1.t3 L1 repetitive element ;, mRNA seq	0.062
	132741	AA394133	Hs.55898	ESTs; Highly similar to OASIS protein [M.musculus]	0.062
50	134437	M26041	Hs.198253	major histocompatibility complex; class II; DQ alpha 1	0.062
	107664	AA010594	Hs.5326	ESTs; Moderately similar to pim-1 protein [H.sapiens]	0.062
	120844	AA349417	Hs.96917	ESTs	0.062
	101574	M34182	Hs.158029	protein kinase; cAMP-dependent; catalytic; gamma	0.062
	131219	C00476	Hs.24395	small inducible cytokine subfamily B (Cys-X-Cys); member 14 (BRAK)	0.062
55	103495	Y09022	Hs.153591	Not56 (D. melanogaster)-like protein	0.062
	129607	AA404594	Hs.11607	ESTs	0.062
	106467	AA450040	Hs.154162	ADP-ribosylation factor-like 2	0.062
	128841	T16358	Hs.106443	ESTs	0.062
	100515	HG1723-HT1729		Macrophage Scavenger Receptor, Alt. Splice 2	0.062
60	119332	T54095		ESTs; Weakly similar to !! ALU SUBFAMILY J WARNING ENTRY !! [H.sapiens]	0.062
	134516	AA171939	Hs.23413	ESTs	0.062
	135012	X73608	Hs.93029	sparc/osteonectin; cwcv and kazal-like domains proteoglycan (testican)	0.063
	103575	Z26256		H.sapiens isoform 1 gene for L-type calcium channel, exon 1	0.063
	115514	AA297739	Hs.55609	ESTs; Weakly similar to ISOLEUCYL-TRNA SYNTHETASE; CYTOPLASMIC [H.sapiens]	0.063
65	103996	AA321355		EST2393 Bone marrow Homo sapiens cDNA 5' end, mRNA sequence	0.063
	110505	H55992	Hs.20495	DKFZP434F011 protein	0.063
	133912	X62744	Hs.77522	major histocompatibility complex; class II; DM alpha	0.063
	129581	M33600	Hs.180255	major histocompatibility complex; class II; DR beta 1	0.063

130139	R38260	Hs.150922	BCS1 (yeast homolog)-like	0.064
105817	AA397825	Hs.5307	synaptopodin	0.064
134658	AA410617	Hs.178009	ESTs	0.064
100306	D50495	Hs.80598	transcription elongation factor A (SII); 2	0.064
100277	D42053	Hs.75890	site-1 protease (subtilisin-like; sterol-regulated; cleaves sterol regulatory element binding proteins)	0.064
133116	D61259	Hs.6529	ESTs	0.064
134909	AA521488	Hs.90998	KIAA0128 protein	0.064
130319	X74794	Hs.154443	minichromosome maintenance deficient (S. cerevisiae) 4	0.064
132057	AA102489	Hs.173484	ESTs	0.064
108334	AA070473		zm7c8.s1 Stratagene neuroepithelium (#937231) Homo sapiens cDNA clone IMAGE:5399 3', mRNA sequence	0.064
129763	F10815	Hs.12373	KIAA0422 protein	0.064
135112	T67464	Hs.94617	ESTs; Weakly similar to predicted using Genefinder [C.elegans]	0.064
122269	AA436856	Hs.98910	ESTs	0.064
133082	AA457129	Hs.6455	RuvB (E coli homolog)-like 2	0.064
113213	T58607		ya94a02.s1 Stratagene placenta (#937225) Homo sapiens cDNA clone IMAGE:69290 3', mRNA sequence.	0.065
106228	AA429290	Hs.17719	ESTs	0.065
130192	Y12661	Hs.171014	VGF nerve growth factor inducible	0.065
104894	AA054087	Hs.18858	phospholipase A2; group IVC (cytosolic; calcium-independent)	0.065
103508	Y10141		H.sapiens DAT1 gene, partial, VNTR	0.065
128474	U40671	Hs.100299	ligase III; DNA; ATP-dependent	0.065
134012	AA417821	Hs.237924	ESTs; Highly similar to CGI-69 protein [H.sapiens]	0.065
134536	AA457735	Hs.850	IMP (inosine monophosphate) dehydrogenase 1	0.065
111714	R23146	Hs.23466	ESTs	0.065
110521	H57060	Hs.108268	ESTs	0.065
103282	X80198	Hs.77628	steroidogenic acute regulatory protein related	0.065
113921	W80730	Hs.28355	ESTs	0.065
129331	N93465	Hs.110453	ESTs; Highly similar to CGI-38 protein [H.sapiens]	0.065
111316	N74597	Hs.180535	ESTs; Weakly similar to mitogen inducible gene mig-2 [H.sapiens]	0.065
135138	AA036794	Hs.95196	ESTs; Weakly similar to T20B12.3 [C.elegans]	0.065
107289	T10792	Hs.172098	ESTs	0.065
121405	AA406083	Hs.98007	ESTs	0.065
124965	T16275	Hs.106359	ESTs	0.065
106595	AA456933	Hs.174481	ESTs	0.066
100106	AF015910		Homo sapiens unknown protein mRNA, partial cds	0.066
134715	AA282757	Hs.89040	preprononiceptin	0.066
135367	AA480109	Hs.9963	TYRO protein tyrosine kinase binding protein	0.066
111533	R08548	Hs.251651	EST	0.066
128509	R53109	Hs.247362	dimethylarginine dimethylaminohydrolase 2	0.066
101030	J05037	Hs.76751	serine dehydratase	0.066
102753	U80226		Human gamma-aminobutyric acid transaminase mRNA, partial cds	0.067
126991	R31652	Hs.821	biglycan	0.067
109583	F02322	Hs.26135	ESTs	0.067
119241	T12559	Hs.221382	ESTs	0.067
130569	AA156597	Hs.256441	EST; Moderately similar to CGI-136 protein [H.sapiens]	0.067
112926	T10316	Hs.4302	ESTs	0.067
120495	AA256073	Hs.190626	ESTs	0.067
130931	AA278412	Hs.21346	ESTs; Weakly similar to F42C5.7 gene product [C.elegans]	0.067
129982	M87789	Hs.140	immunoglobulin gamma 3 (Gm marker)	0.067
133832	H03387	Hs.241305	estrogen-responsive B box protein	0.067
110697	H93721	Hs.20798	ESTs	0.067
121183	AA400138	Hs.97703	ESTs	0.067
130953	U12707	Hs.2157	Wiskott-Aldrich syndrome (ecezema-thrombocytopenia)	0.067
102218	U24183	Hs.75160	phosphofructokinase; muscle	0.067
114181	Z39079	Hs.8021	KIAA1058 protein	0.067
116581	D51267	Hs.82148	ribosomal protein S12	0.067
132498	T87708	Hs.50098	ESTs	0.068
103788	AA096014	Hs.9527	ESTs; Highly similar to HSPC013 [H.sapiens]	0.068
102459	U48936		Human amiloride-sensitive epithelial sodium channel gamma subunit mRNA, 5' end, partial cds	0.068
100373	D79999	Hs.77225	ADP-ribosyltransferase (NAD+; poly (ADP-ribose) polymerase)-like 1	0.068
132717	AA203321	Hs.151696	DKFZP727G051 protein	0.068
128863	D87462	Hs.106674	BRCA1 associated protein-1 (ubiquitin carboxy-terminal hydrolase)	0.068
115193	AA262029	Hs.88218	ESTs	0.068
124558	N66046	Hs.141605	ESTs	0.069
117225	N20392	Hs.42846	ESTs	0.069
110665	H83380	Hs.32757	ESTs	0.069

132905	U70663	Hs.182965	Kruppel-like factor 4 (gut)	0.069
105778	AA348910	Hs.153299	DOM-3 (C. elegans) homolog Z	0.069
134770	R72079	Hs.89575	CD79B antigen (immunoglobulin-associated beta)	0.069
123097	AA485869	Hs.105671	ESTs	0.069
100750	HG3523-HT4899		Proto-Oncogene C-Myc, Alt. Splice 3, Orf 114	0.069
125091	T91518		ye20f05.s1 Stratagene lung (#937210) H sapiens cDNA clone IMAGE: 3' similar to contains Alu repetitive element;contains MER12 repetitive element; mRNA sequence.	0.069
100756	HG3565-HT3768		Zinc Finger Protein (Gb:M88357)	0.069
113483	T87768	Hs.16439	ESTs	0.069
101119	L09708	Hs.2253	complement component 2	0.069
102286	U31628	Hs.12503	interleukin 15 receptor; alpha	0.07
135349	D83174	Hs.9930	collagen-binding protein 2 (colligen 2)	0.07
100991	J03764	Hs.82085	plasminogen activator inhibitor; type I	0.07
133675	AA443720	Hs.7551	ESTs; Weakly similar to T25G3.1 [C.elegans]	0.07
105422	AA251014	Hs.12210	ESTs	0.07
102932	X13334	Hs.75627	CD14 antigen	0.07
119147	R58878	Hs.65739	ESTs	0.07
104900	AA055048	Hs.180481	ESTs; Weakly similar to ACROSIN PRECURSOR [H.sapiens]	0.07
133185	AA481404	Hs.6686	ESTs	0.07
115496	AA290674	Hs.71819	eukaryotic translation initiation factor 4E binding protein 1	0.07
121005	AA398332	Hs.97613	ESTs	0.07
124869	R69098	Hs.28728	ESTs; Weakly similar to F55A12.9 [C.elegans]	0.071
129154	N23673	Hs.108969	mannosidase; alpha; class 2B; member 1	0.071
112161	R48295		ESTs; Wkly smlr to !! ALU SUBFAMILY J WARNING ENTRY !! [H.sapiens]	0.071
125251	W87486	Hs.141464	ESTs	0.071
134298	J00116	Hs.81343	collagen; type II; alpha 1 (primary osteoarthritis; spondyloepiphyseal dysplasia; congenital)	0.071
119745	W70264	Hs.58093	ESTs	0.071
131306	AA232686	Hs.25489	ESTs	0.071
107776	AA018820	Hs.221147	ESTs	0.071
134271	AA199630	Hs.184456	ESTs; Wkly smlr to !! ALU SUBFAMILY SX WARNING ENTRY !! [H.sapiens]	0.071
101798	M85220		Accession not listed in Genbank	0.071
135402	S76942	Hs.99922	dopamine receptor D4	0.071
118742	N74052	Hs.50424	EST	0.071
131867	N64656	Hs.3353	Homo sapiens clone 24940 mRNA sequence	0.071
102923	X12517	Hs.1063	small nuclear ribonucleoprotein polypeptide C	0.072
100775	HG371-HT26388		Mucin 1, Epithelial, Alt. Splice 9	0.072
111020	N54361	Hs.185726	ESTs	0.072
134224	X80822	Hs.163593	ribosomal protein L18a	0.072
124059	F13673	Hs.99769	ESTs	0.072
133972	AA160743	Hs.78019	Homo sapiens clone 24432 mRNA sequence	0.072
129681	AA436009	Hs.178186	ESTs; Weakly similar to WASP-family protein [H.sapiens]	0.072
103065	X58399	Hs.81221	Human L2-9 transcript of unrearranged immunoglobulin V(H)5 pseudogene	0.072
124966	T19271	Hs.155560	calnexin	0.072
112270	R53021	Hs.203358	ESTs	0.072
116704	F10183	Hs.66140	EST	0.072
129890	M13699	Hs.111461	ceruloplasmin (ferroxidase)	0.072
127345	AA972008	Hs.166253	ESTs; Highly similar to KIAA0476 protein [H.sapiens]	0.072
112436	R63090	Hs.28391	ESTs	0.072
114531	AA053033	Hs.203330	ESTs	0.072
135122	H99080	Hs.94814	ESTs	0.072
103934	AA281338	Hs.134200	Homo sapiens mRNA; cDNA DKFZp564C186 (from clone DKFZp564C186)	0.072
109363	AA215369	Hs.185764	ESTs; Weakly similar to hypothetical protein [H.sapiens]	0.072
112647	R83329	Hs.33403	ESTs	0.073
127083	Z44079	Hs.91608	otoferlin	0.073
133027	AA402624	Hs.63236	synuclein; gamma (breast cancer-specific protein 1)	0.073
122086	AA432121	Hs.250986	EST	0.073
110405	H47542	Hs.33962	ESTs	0.073
128697	AB002344	Hs.103915	KIAA0346 protein	0.073
112221	R50380	Hs.25670	ESTs	0.073
100478	HG1067-HT1067		Mucin (Gb:M22406)	0.073
115598	AA400129	Hs.65735	ESTs	0.073
132491	AA227137	Hs.4984	KIAA0828 protein	0.073
101655	M60299		Human alpha-1 collagen type II gene, exons 1, 2 and 3	0.073
106018	AA411887	Hs.34737	ESTs	0.073
129683	W05348	Hs.158196	DKFZP434B103 protein	0.073
134137	F10045	Hs.79347	KIAA0211 gene product	0.073
114008	W89128	Hs.19872	ESTs	0.073